

### **Protein Sequence Searches - 10/8/04**

All of the sequence databases on the ABSS have been updated. A change has occurred in the protein databases.

- Two protein databases, SPTREMBL and SwissProt, are now produced as a single, merged database called UniProt.
- Results from UniProt have the file extension **.rup**.
- Sequences in UniProt are identified by the same ID that had been used in SPTREMBL or SwissProt.
- In instances where the database curators have determined that an SPTREMBL record and a SwissProt record represent the same sequence, the two records have been merged into one. Both IDs are present in the record. Any differences found between the two sequences are recorded in the FT (feature table) fields.

If you have any questions regarding these changes or your results, please contact any STIC searcher.

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 3, 2004, 11:25:27 ; Search time 2090.5 Seconds

(without alignments)  
135.727 Million cell updates/sec

Title: US-10-030-194A-6

Perfect score: 24

Sequence: 1 GYXVEX 6

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool.p/US10030194/runat.01112004.184847.28955/app.query.fasta\_1.398  
-DB=GenEmbl -QWTF=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000  
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=75 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-NO.MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sv.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	91.7	16	6 AR062548	AR062548 Sequence
C 2	22	91.7	21	6 AR307356	AR307356 Sequence
C 3	22	91.7	21	6 AR307357	AR307357 Sequence
C 4	22	91.7	21	6 BD070798	BD070798 Method to

C	5	22	91.7	21	6	BD070800	Method to
	6	22	91.7	28	6	AX135256	Sequence
	7	22	91.7	31	6	AX554130	Sequence
	8	22	91.7	35	6	AX554131	Sequence
C	9	22	91.7	48	6	AR075824	Sequence
C	10	22	91.7	48	6	E30457	E30457 Method for
	11	22	91.7	60	6	CQ544105	Sequence
C	12	22	91.7	60	6	CQ552569	Sequence
	13	22	91.7	65	6	CQ533924	Sequence
C	14	22	91.7	75	6	AR357220	Sequence
	15	22	91.7	95	6	AR035489	Sequence
	16	22	91.7	100	6	CQ001247	Sequence
	17	22	91.7	100	6	CQ001248	Sequence
	18	22	91.7	100	6	CQ001249	Sequence
	19	22	91.7	100	6	AR366373	Sequence
	20	22	91.7	100	6	AR366376	Sequence
	21	22	91.7	100	6	AR435569	Sequence
	22	22	91.7	100	6	AR435572	Sequence
	23	22	91.7	108	6	AX088741	Sequence
C	24	22	91.7	127	6	CQ713868	Sequence
	25	22	91.7	129	6	CQ672500	Sequence
C	26	22	91.7	136	6	CQ113946	Sequence
	27	22	91.7	136	6	CQ152829	Sequence
C	28	22	91.7	136	6	CQ236146	Sequence
	29	22	91.7	136	6	CQ273724	Sequence
C	30	22	91.7	136	6	CQ348024	Sequence
	31	22	91.7	152	6	CQ832356	Sequence
C	32	22	91.7	161	10	F369389S54	Sequence
	33	22	91.7	165	11	AX663632	Sequence
C	34	22	91.7	166	6	AR358585	Sequence
	35	22	91.7	169	6	CQ434907	Sequence
C	36	22	91.7	172	9	HS301D7	Sequence
	37	22	91.7	172	11	AX65325	Sequence
C	38	22	91.7	172	11	AX546436	Sequence
	39	22	91.7	175	11	AX465370	Sequence
C	40	22	91.7	176	11	AX465345	Sequence
	41	22	91.7	184	11	AX831542	Sequence
C	42	22	91.7	184	11	AX664061	Sequence
	43	22	91.7	186	8	BT011699	Sequence
C	44	22	91.7	188	9	HUMINSRA02	Sequence
	45	22	91.7	189	6	CQ059104	Sequence
C	46	22	91.7	189	6	CQ183536	Sequence
	47	22	91.7	189	6	CQ269587	Sequence
C	48	22	91.7	189	6	CQ306668	Sequence
	49	22	91.7	193	6	AR206039	Sequence
C	50	22	91.7	194	9	HS39C5F	Sequence
	51	22	91.7	200	6	CQ713883	Sequence
	52	22	91.7	201	11	BV175451	Sequence
	53	22	91.7	201	11	BV204514	Sequence
C	54	22	91.7	202	6	AR269722	Sequence
	55	22	91.7	204	11	AX284270	Sequence
	56	22	91.7	206	6	AX919636	Sequence
	57	22	91.7	206	6	BD055169	Sequence
	58	22	91.7	220	12	SYNDHFR11	Sequence
	59	22	91.7	221	9	F260679801	Sequence
C	60	22	91.7	225	6	AR396502	Sequence
	61	22	91.7	228	6	AR395391	Sequence
C	62	22	91.7	237	11	AX323626	Sequence
	63	22	91.7	237	11	G02926	Sequence
C	64	22	91.7	238	11	G28679	Sequence
	65	22	91.7	240	4	AF483760	Sequence
	66	22	91.7	241	3	AF549480	Sequence
C	67	22	91.7	242	11	AV088606	Sequence
	68	22	91.7	243	11	AL823809	Sequence
C	69	22	91.7	245	11	AX548089	Sequence
	70	22	91.7	247	6	AX694666	Sequence
	71	22	91.7	248	11	HSB0192H9	Sequence
	72	22	91.7	250	9	HS184G6R	Sequence
C	73	22	91.7	251	6	AX188844	Sequence
	74	22	91.7	258	11	G81230	Sequence
C	75	22	91.7	260	1	BSPKJ17	Sequence
	76	22	91.7	265	5	AF145907	Sequence
C	77	22	91.7	268	1	AF406353	Sequence

BD070800	Method to
AX135256	Sequence
AX554130	Sequence
AX554131	Sequence
AR075824	Sequence
E30457	Method for
CQ544105	Sequence
CQ552569	Sequence
CQ533924	Sequence
AR357220	Sequence
AR035489	Sequence
CQ001247	Sequence
CQ001248	Sequence
CQ001249	Sequence
AR366373	Sequence
AR366376	Sequence
AR435569	Sequence
AR435572	Sequence
AX088741	Sequence
CQ713868	Sequence
CQ672500	Sequence
CQ113946	Sequence
CQ152829	Sequence
CQ236146	Sequence
CQ273724	Sequence
CQ348024	Sequence
CQ832356	Sequence
AX663632	Sequence
AR358585	Sequence
CQ434907	Sequence
Z79819	H.sapiens c
AX65325	Arabidops
AX546436	Arabidops
AX465370	Arabidops
AX465345	Arabidops
AX831542	Arabidops
AX664061	Arabidops
BT011699	Arabidops
M27197	Human insul
CQ059104	Sequence
CQ183536	Sequence
CQ269587	Sequence
CQ306668	Sequence
AR206039	Sequence
Z55399	H.sapiens c
CQ713883	Sequence
CV175451	sqnm79864
CV204514	sqnm21547
AR269722	Sequence
AX919636	Sequence
BD055169	Sequence
M20407	Synthetic m
AF260879	Homo sapi
AR396502	Sequence
AR395391	Sequence
AX323626	Arabidops
G02926	human STS W
G28679	sqm3790 Er
AF483760	Bos tauru
AF549480	Coccotryp
AV088606	RPAMSEFQ
AL823809	Arabidops
AX548089	Arabidops
AX694666	Sequence
Z53115	H.sapiens (
Z57571	H.sapiens c
AX188844	Sequence
G81230	g210P6001RD
Z83249	Rhizobiacea
AF145907	Columba i
AF406353	Unculture

c 78	22	91.7	271	5	AF145900	AF145900	Columba i	151	22	91.7	349	3	AF241047S1	AF241047 Tachyuroop
79	22	91.7	271	11	AU025999	AU025999	Rattus no	152	22	91.7	351	14	AF534999	AF534999 Hepatitis
80	22	91.7	273	1	HEAPSHPROF	D28887	Haemophilus	c 153	22	91.7	352	5	AF145911	AF145911 Columba i
81	22	91.7	273	6	AR236625	AR236625	Sequence	c 154	22	91.7	353	5	AF145904	AF145904 Columba i
82	22	91.7	273	6	AX127450	AX127450	Sequence	c 155	22	91.7	353	6	AX071045	AX071045 Sequence
83	22	91.7	273	9	HGJ005385	AJ005385	Homo sapi	156	22	91.7	353	12	SYNDHFRS	K02125 Plasmid R67
84	22	91.7	277	6	AX912140	AX912140	Sequence	157	22	91.7	354	6	CO688370	CO688370 Sequence
85	22	91.7	277	6	BD047673	BD047673	Sequence	c 158	22	91.7	354	11	BX546170	BX546170 Arabidops
86	22	91.7	278	5	AF145899	AF145899	Columba i	c 159	22	91.7	357	8	AX507027	AX507027 Sequence
87	22	91.7	283	5	AX380786	AX380786	Meleagris	c 160	22	91.7	357	14	AX091677	AX091677 Arabidops
88	22	91.7	285	6	AX910804	AX910804	Sequence	c 161	22	91.7	357	14	AX486250	AX486250 Hepatitis
89	22	91.7	285	6	BD046337	BD046337	Sequence	c 162	22	91.7	358	11	BX546198	BX546198 Arabidops
c 90	22	91.7	286	3	AY466349	AY466349	Litopenae	c 163	22	91.7	359	5	AF145918	AF145918 Unculture
91	22	91.7	286	6	AS1490	AS1490	Sequence	c 164	22	91.7	360	6	CO051298	CO051298 Columba i
92	22	91.7	286	6	AK062544	AK062544	Sequence	166	22	91.7	360	6	CO066348	CO066348 Sequence
93	22	91.7	286	6	AK062545	AK062545	Sequence	c 165	22	91.7	360	6	CO093394	CO093394 Sequence
c 94	22	91.7	286	6	AX188845	AX188845	Sequence	167	22	91.7	360	6	CO093394	CO093394 Sequence
95	22	91.7	287	6	AX114606	AX114606	Sequence	168	22	91.7	360	6	CO132171	CO132171 Sequence
96	22	91.7	287	6	AX114732	AX114732	Sequence	169	22	91.7	360	6	CO170738	CO170738 Sequence
97	22	91.7	287	6	AX118662	AX118662	Sequence	170	22	91.7	360	6	CO199887	CO199887 Sequence
98	22	91.7	287	6	AX128079	AX128079	Sequence	171	22	91.7	360	6	CO215357	CO215357 Sequence
99	22	91.7	287	6	AX644680	AX644680	Sequence	172	22	91.7	360	6	CO253965	CO253965 Sequence
100	22	91.7	287	9	HUMINSR18	M32839	Human insul	173	22	91.7	360	6	CO328087	CO328087 Sequence
101	22	91.7	289	6	CO701125	CO701125	Sequence	c 174	22	91.7	361	1	AY476766	AY476766 Unculture
c 102	22	91.7	291	5	AF145909	AF145909	Columba i	c 175	22	91.7	362	5	AF145916	AF145916 Columba i
c 103	22	91.7	291	11	AF639419	BX664446	Arabidops	c 176	22	91.7	362	11	BV150905	BV150905 P2A02092-
c 104	22	91.7	292	1	R67DHR	AF394119	Unculture	c 177	22	91.7	362	5	AF145919	AF145919 Columba i
c 105	22	91.7	296	11	BX284286	K02118	Plasmid R67	c 178	22	91.7	364	5	AF145920	AF145920 Columba i
c 106	22	91.7	296	11	G64955	BX664446	Arabidops	c 179	22	91.7	364	11	AL823751	AL823751 Arabidops
c 107	22	91.7	296	11	AF276609	BX284286	Arabidops	c 180	22	91.7	365	11	AL823501	AL823501 Arabidops
c 108	22	91.7	298	14	AF174630	G64955	G-104334 Ra	c 181	22	91.7	366	6	BD242121	BD242121 Compounds
c 109	22	91.7	298	1	PSP404556	AF174630	Unculture	c 182	22	91.7	366	6	AR237304	AR237304 Sequence
c 110	22	91.7	298	1	ATCMTCC	AF174630	Unculture	c 183	22	91.7	366	6	AR278328	AR278328 Sequence
c 111	22	91.7	299	5	Q744101	AJ404556	Unculture	c 184	22	91.7	366	6	AR367024	AR367024 Sequence
c 112	22	91.7	300	6	CTC44101	M64890	Attractoste	c 185	22	91.7	366	6	AR370920	AR370920 Sequence
c 113	22	91.7	300	6	BD128788	CO744101	Sequence	c 186	22	91.7	366	6	AR392425	AR392425 Sequence
c 114	22	91.7	303	1	UAU78912	BD128788	Human gen	c 187	22	91.7	366	6	AR400060	AR400060 Sequence
c 115	22	91.7	305	6	CO419065	Y07925	M.antiirrhin	c 188	22	91.7	366	6	AR405327	AR405327 Sequence
c 116	22	91.7	305	6	AX308990	Y07925	Unidentiife	c 189	22	91.7	366	6	AR439531	AR439531 Sequence
c 117	22	91.7	314	6	AR134682	CO419065	Sequence	c 190	22	91.7	366	6	AX106434	AX106434 Sequence
c 118	22	91.7	314	6	AR152043	AR134682	Sequence	c 191	22	91.7	366	6	AX140725	AX140725 Sequence
c 119	22	91.7	314	6	AR242268	AR152043	Sequence	c 192	22	91.7	366	6	AX200585	AX200585 Sequence
c 120	22	91.7	318	6	AX310498	AR242268	Sequence	c 193	22	91.7	366	6	AX267241	AX267241 Sequence
c 121	22	91.7	319	5	AF145917	AX310498	Sequence	c 194	22	91.7	366	6	AX608656	AX608656 Sequence
c 122	22	91.7	322	6	BD232190	BD232190	Novel ser	c 195	22	91.7	366	6	BD070357	BD070357 Compounds
c 123	22	91.7	322	6	AR366786	AR366786	Sequence	c 196	22	91.7	368	1	AF432286	AF432286 Unculture
c 124	22	91.7	326	6	AX155253	AX155253	Sequence	c 197	22	91.7	370	11	BV150907	BV150907 P2A02092-
c 125	22	91.7	326	6	AX358332	AX358332	Sequence	c 198	22	91.7	372	6	AR051468	AR051468 Sequence
c 126	22	91.7	327	5	AF145901	AX358332	Sequence	c 199	22	91.7	372	6	AR072608	AR072608 Sequence
c 127	22	91.7	327	8	AY558383	AF145901	Columba i	c 200	22	91.7	372	6	AR073153	AR073153 Sequence
c 128	22	91.7	339	11	G04263	AY558383	Saccharom	c 201	22	91.7	373	3	MCEINT1	Y07926 M.certus ri
c 129	22	91.7	330	3	MANITS1	G04263	human STS W	c 202	22	91.7	374	11	G31669	G31669 sWS1082 Er
c 130	22	91.7	330	5	AF145902	Y07925	M.antiirrhin	c 203	22	91.7	375	3	AF169906	AF169906 Eueides i
c 131	22	91.7	330	6	E51213	AF145902	Columba i	c 204	22	91.7	375	3	AY090515	AY090515 Hoplia ar
c 132	22	91.7	330	6	I08256	E51213	Disease tol	c 205	22	91.7	378	6	A00040	A00040 P.chrysogen
c 133	22	91.7	330	8	AB022172	I08256	Sequence	c 206	22	91.7	378	6	A12434	A12434 EcoRI casse
c 134	22	91.7	332	5	AF145903	AB022172	Oryza sat	c 207	22	91.7	380	1	AF007255	AF007255 Unidentif
c 135	22	91.7	332	5	AF145905	AF145903	Columba i	c 208	22	91.7	380	1	AF007258	AF007258 Unidentif
c 136	22	91.7	332	5	AF145908	AF145905	Columba i	c 209	22	91.7	381	8	BT012085	BT012085 Arabidops
c 137	22	91.7	333	5	AF145912	AF145908	Columba i	c 210	22	91.7	382	1	AF078189	AF078189 Grassland
c 138	22	91.7	333	9	HUMPPHB	AF145912	Columba i	c 211	22	91.7	384	6	AR274015	AR274015 Sequence
c 139	22	91.7	335	11	BX284706	M17257	Human/papil	c 212	22	91.7	384	6	AR277596	AR277596 Sequence
c 140	22	91.7	336	8	CNS019WN	BX284706	Arabidops	c 213	22	91.7	384	6	AR441721	AR441721 Sequence
c 141	22	91.7	336	8	CNS01BI7	AL112191	Botrytis	c 214	22	91.7	384	6	AX208311	AX208311 Sequence
c 142	22	91.7	337	5	AF145913	AL112191	Botrytis	c 215	22	91.7	384	6	AX369058	AX369058 Sequence
c 143	22	91.7	339	6	CO526947	AL14263	Botrytis	c 216	22	91.7	384	8	BT011713	BT011713 Arabidops
c 144	22	91.7	342	11	AL807279	AF145913	Columba i	c 217	22	91.7	384	11	BV093373	BV093373 RPAMSEQ0
c 145	22	91.7	342	6	AR135863	CO526947	Sequence	c 218	22	91.7	385	6	AX262127	AX262127 Sequence
c 146	22	91.7	342	6	CO434613	AR135863	Sequence	c 219	22	91.7	385	8	AF263300	AF263300 Sphaerops
c 147	22	91.7	342	6	CO434613	AR135985	Sequence	c 220	22	91.7	385	8	AF010286	AF010286 Unidentif
c 148	22	91.7	342	5	AF145910	AR395392	Sequence	c 221	22	91.7	387	1	AF010286	AF010286 Unidentif
c 149	22	91.7	345	5	AF145910	AF145910	Columba i	c 222	22	91.7	387	1	UPR431349	AJ431349 Unculture
c 150	22	91.7	345	6	AR396783	AF145910	Columba i	c 223	22	91.7	387	6	I08257	I08257 Sequence 3

C 224	22	91.7	388	1	AF268844	Unculture	AF268844	Unculture	C 297	22	91.7	428	5	AF113628	Elseya ge
C 225	22	91.7	388	11	AF268851	Unculture	AF268851	Unculture	C 298	22	91.7	428	6	CQ518442	Sequence
C 226	22	91.7	388	11	G79421	Sequence	G79421	S210P6193RA	C 299	22	91.7	430	1	AF460890	Marine ba
C 227	22	91.7	390	1	AF010285	Unculture	AF010285	Unculture	C 300	22	91.7	431	6	CQ505357	Sequence
C 228	22	91.7	392	1	AV344910	Sequence	AV344910	Unculture	C 301	22	91.7	431	6	CQ516909	Sequence
C 229	22	91.7	393	6	AX261950	Sequence	AX261950	Sequence	C 302	22	91.7	431	6	BD058381	Secreted
C 230	22	91.7	394	1	AB022479	Unculture	AB022479	Unculture	C 303	22	91.7	432	1	AF022396	Uncidentif
C 231	22	91.7	395	1	AX344893	Unculture	AX344893	Unculture	C 304	22	91.7	432	1	AY476787	Unculture
C 232	22	91.7	395	6	CQ069637	Sequence	CQ069637	Sequence	C 305	22	91.7	432	1	AY476798	Unculture
C 233	22	91.7	395	6	CQ096749	Sequence	CQ096749	Sequence	C 306	22	91.7	433	1	AF141544	Unculture
C 234	22	91.7	395	6	CQ133546	Sequence	CQ133546	Sequence	C 307	22	91.7	433	1	AF141545	Unculture
C 235	22	91.7	395	6	CQ173979	Sequence	CQ173979	Sequence	C 308	22	91.7	433	1	AY476763	Unculture
C 236	22	91.7	395	6	CQ218840	Sequence	CQ218840	Sequence	C 309	22	91.7	434	1	AY476796	Unculture
C 237	22	91.7	395	6	CQ257421	Sequence	CQ257421	Sequence	C 310	22	91.7	434	1	RSU64002	U64002 Rhizobium s
C 238	22	91.7	395	6	CQ294646	Sequence	CQ294646	Sequence	C 311	22	91.7	435	6	CQ751262	Sequence
C 239	22	91.7	395	6	CQ331309	Sequence	CQ331309	Sequence	C 312	22	91.7	436	1	AF439428	Unculture
C 240	22	91.7	395	6	CQ445143	Sequence	CQ445143	Sequence	C 313	22	91.7	436	6	CQ396232	Sequence
C 241	22	91.7	395	11	BV150899	Sequence	BV150899	PZA02092-	C 314	22	91.7	436	6	CQ402553	Sequence
C 242	22	91.7	396	1	AY476765	Unculture	AY476765	Unculture	C 315	22	91.7	437	1	AF547428	Unculture
C 243	22	91.7	396	6	AX208676	Sequence	AX208676	Sequence	C 316	22	91.7	437	6	CQ396047	Sequence
C 244	22	91.7	396	10	AB013468	Mus muscu	AB013468	Mus muscu	C 317	22	91.7	437	6	CQ402373	Sequence
C 245	22	91.7	396	11	AL831529	Arabidops	AL831529	Arabidops	C 318	22	91.7	438	1	AF365448	Unculture
C 246	22	91.7	398	1	AF078274	Grassland	AF078274	Grassland	C 319	22	91.7	439	3	AB110720	Dorcus re
C 247	22	91.7	398	11	BF150902	Sequence	BV150902	PZA02092-	C 320	22	91.7	439	3	AB110721	Dorcus re
C 248	22	91.7	400	1	AY149784	Unculture	AY149784	Unculture	C 321	22	91.7	441	1	AF439425	Unculture
C 249	22	91.7	400	6	CQ475389	Sequence	CQ475389	Sequence	C 322	22	91.7	441	1	AF547399	Unculture
C 250	22	91.7	400	11	G17749	human STS S	G17749	human STS S	C 323	22	91.7	442	1	PSP309981	Paracoccu
C 251	22	91.7	401	11	BV189181	Sequence	BV189181	sgm16265	C 324	22	91.7	443	11	AL823496	Arabidops
C 252	22	91.7	402	1	AY476772	Unculture	AY476772	Unculture	C 325	22	91.7	443	11	BV093437	RPAMWSEQO
C 253	22	91.7	402	1	AY511690	Unculture	AY511690	Unculture	C 326	22	91.7	445	1	AF439422	Unculture
C 254	22	91.7	402	3	ASP409931	Ammonia s	AY409931	Ammonia s	C 327	22	91.7	445	1	AF439424	Unculture
C 255	22	91.7	402	6	CQ173609	Sequence	CQ173609	Sequence	C 328	22	91.7	445	11	AL823476	Arabidops
C 256	22	91.7	402	6	CQ218369	Sequence	CQ218369	Sequence	C 329	22	91.7	445	11	AL939383	Arabidops
C 257	22	91.7	402	6	CQ330918	Sequence	CQ330918	Sequence	C 330	22	91.7	446	1	AF182018	Sulfitoba
C 258	22	91.7	403	1	AY476774	Unculture	AY476774	Unculture	C 331	22	91.7	446	1	AF361648	Unculture
C 259	22	91.7	404	11	BV150908	Sequence	BV150908	PZA02092-	C 332	22	91.7	446	1	AF365750	Unculture
C 260	22	91.7	405	1	AY476768	Unculture	AY476768	Unculture	C 333	22	91.7	446	1	AF439426	Unculture
C 261	22	91.7	405	11	G23626	human STS W	G23626	human STS W	C 334	22	91.7	447	11	AL845153	Arabidops
C 262	22	91.7	406	3	ABA228541	Ammonia b	AJ228541	Ammonia b	C 335	22	91.7	448	1	AF025323	Roseobact
C 263	22	91.7	406	3	AF337621	Blackburn	AF337621	Blackburn	C 336	22	91.7	448	1	AF365542	Unculture
C 264	22	91.7	406	11	BV150901	Sequence	BV150901	PZA02092-	C 337	22	91.7	450	11	AF235256	Sus scrof
C 265	22	91.7	410	1	AF254099	Alpha pro	AF254099	Alpha pro	C 338	22	91.7	451	1	ZSU63957	U63957 Zoogloea sp
C 266	22	91.7	411	1	AF287021	Unculture	AF287021	Unculture	C 339	22	91.7	451	1	UAU64009	Alpha proce
C 267	22	91.7	411	12	AY657381	Synthetic	AY657381	Synthetic	C 340	22	91.7	451	5	PQNMTCR	M98393 Columba ino
C 268	22	91.7	412	1	ECO429132	Escherich	AJ429132	Escherich	C 341	22	91.7	451	11	AL823529	Arabidops
C 269	22	91.7	414	1	AF439429	Unculture	AF439429	Unculture	C 342	22	91.7	451	11	AL824472	Arabidops
C 270	22	91.7	414	3	ASP228525	Ammonia s	AJ228525	Ammonia s	C 343	22	91.7	452	1	AY476773	Unculture
C 271	22	91.7	414	6	AK384728	Sequence	AK384728	Sequence	C 344	22	91.7	452	6	CQ679902	Sequence
C 272	22	91.7	416	6	CQ394028	Sequence	CQ394028	Sequence	C 345	22	91.7	453	11	BV103843	Arabidops
C 273	22	91.7	416	6	CQ400394	Sequence	CQ400394	Sequence	C 346	22	91.7	455	11	DM7B5S	AF008827 Drosophil
C 274	22	91.7	417	1	AF547398	Unculture	AF547398	Unculture	C 347	22	91.7	456	1	AF522922	Unculture
C 275	22	91.7	417	11	G53531	SHGC-81044	G53531	SHGC-81044	C 348	22	91.7	457	1	AY100684	Alpha pro
C 276	22	91.7	418	1	AF439421	Unculture	AF439421	Unculture	C 349	22	91.7	457	8	AY185355	Brassica
C 277	22	91.7	418	6	AF447881	Sequence	AF447881	Sequence	C 350	22	91.7	458	11	AL823510	Arabidops
C 278	22	91.7	420	1	AB088414	Mesorhizo	AB088414	Mesorhizo	C 351	22	91.7	458	11	G27000	G27000 human STS S
C 279	22	91.7	420	1	AF439427	Unculture	AF439427	Unculture	C 352	22	91.7	459	1	AY123252	Escherich
C 280	22	91.7	420	1	AF547424	Unculture	AF547424	Unculture	C 353	22	91.7	459	6	CQ644510	Sequence
C 281	22	91.7	420	1	AY149762	Unculture	AY149762	Unculture	C 354	22	91.7	459	11	BV003173	S208P6806
C 282	22	91.7	420	6	AK394734	Sequence	AK394734	Sequence	C 355	22	91.7	460	8	ATH553428	Arabidops
C 283	22	91.7	420	8	AF358741	Polyporus	AF358741	Polyporus	C 356	22	91.7	460	11	AL807643	Arabidops
C 284	22	91.7	421	1	AY476771	Unculture	AY476771	Unculture	C 357	22	91.7	460	11	AL823843	Arabidops
C 285	22	91.7	422	1	AY476739	Unculture	AY476739	Unculture	C 358	22	91.7	460	11	G31452	human STS S
C 286	22	91.7	424	1	AY476739	Unculture	AY476739	Unculture	C 359	22	91.7	461	11	AL823819	Arabidops
C 287	22	91.7	424	6	AK425968	Sequence	AK425968	Sequence	C 360	22	91.7	461	11	BX322120	Arabidops
C 288	22	91.7	424	6	AK986662	Sequence	AK986662	Sequence	C 361	22	91.7	462	10	MUSICKPALI	Mus musculu
C 289	22	91.7	424	6	BD121521	EST and e	BD121521	EST and e	C 362	22	91.7	463	6	CQ457516	Sequence
C 290	22	91.7	425	1	AY476769	Unculture	AY476769	Unculture	C 363	22	91.7	465	1	AF439412	Unculture
C 291	22	91.7	426	1	APR315684	Alpha pro	AJ315684	Alpha pro	C 364	22	91.7	465	1	AF499418	Unculture
C 292	22	91.7	427	1	URH548904	unculture	AJ548904	unculture	C 365	22	91.7	465	11	BV075415	S208P6705
C 293	22	91.7	427	6	AP418876	Sequence	AP418876	Sequence	C 366	22	91.7	466	1	AY476788	Unculture
C 294	22	91.7	427	6	AK979570	Sequence	AK979570	Sequence	C 367	22	91.7	466	1	AY476790	Unculture
C 295	22	91.7	427	6	BD114429	EST and e	BD114429	EST and e	C 368	22	91.7	466	11	AL824433	Arabidops
C 296	22	91.7	428	1	UAL561156	Unculture	AJ561156	Unculture	C 369	22	91.7	467	6	AX401427	Sequence

C 370	22	91.7	467	11	AL824327	AL824327 Arabidops	C 443	22	91.7	511	4	BBA431653	AJ431653 Barbastel
C 371	22	91.7	468	1	AF439420	AF439420 Unculture	444	22	91.7	513	6	AR415506	AR415506 Sequence
C 372	22	91.7	468	11	AL823793	AL823793 Arabidops	445	22	91.7	513	6	AX972340	AX972340 Sequence
C 373	22	91.7	468	14	AY370769	AY370769 Newcastle	446	22	91.7	513	6	BD111059	BD111059 EST and e
C 374	22	91.7	469	6	CQ686457	CQ686457 Sequence	C 447	22	91.7	514	1	AR260287	AR260287 Rhizobium
C 375	22	91.7	470	1	AF4339415	AF4339415 Unculture	448	22	91.7	515	6	AR425094	AR425094 Sequence
C 376	22	91.7	470	1	AF547425	AF547425 Unculture	449	22	91.7	515	6	AX985788	AX985788 Sequence
C 377	22	91.7	470	6	CQ423711	CQ423711 Sequence	450	22	91.7	515	6	BD120647	BD120647 EST and e
C 378	22	91.7	471	6	AF4339410	AF4339410 Unculture	C 451	22	91.7	516	14	AY039379	AY039379 Earthworm
C 379	22	91.7	471	6	CQ745199	CQ745199 Sequence	C 452	22	91.7	516	14	AF108153	AF108153 Human rhl
C 380	22	91.7	472	6	AR357058	AR357058 Sequence	453	22	91.7	519	6	BD232189	BD232189 Novel ser
C 381	22	91.7	472	11	AL823810	AL823810 Arabidops	C 454	22	91.7	521	6	AR366785	AR366785 Sequence
C 382	22	91.7	475	6	AR414421	AR414421 Sequence	455	22	91.7	521	6	AX867799	AX867799 Sequence
C 383	22	91.7	475	6	AX971255	AX971255 Sequence	C 456	22	91.7	525	6	BD147861	BD147861 Primer fo
C 384	22	91.7	475	6	BD109974	BD109974 EST and e	457	22	91.7	525	6	BD147861	BD147861 Primer fo
C 385	22	91.7	476	6	CQ054127	CQ054127 Sequence	458	22	91.7	526	11	GS7491	GS7491 SHGC-103350
C 386	22	91.7	476	6	CQ173666	CQ173666 Sequence	C 459	22	91.7	528	1	AY090122	AY090122 Unculture
C 387	22	91.7	476	6	CQ257025	CQ257025 Sequence	C 460	22	91.7	529	8	AY354280	AY354280 Umbelopsi
C 388	22	91.7	476	6	CQ294167	CQ294167 Sequence	C 461	22	91.7	530	1	UBA419053	AJ419053 unculture
C 389	22	91.7	477	1	AF210717	AF210717 Unculture	462	22	91.7	535	6	AX432800	AX432800 Sequence
C 390	22	91.7	477	1	AF210721	AF210721 Unculture	C 463	22	91.7	535	11	CR384501	CR384501 Arabidops
C 391	22	91.7	477	12	AY657786	AY657786 Synthetic	C 464	22	91.7	536	1	AY476756	AY476756 Unculture
C 392	22	91.7	478	1	AY476791	AY476791 Unculture	465	22	91.7	536	4	PTSMTRRZD	M81139 Pteropus ca
C 393	22	91.7	479	1	AY037757	AY037757 Unculture	C 466	22	91.7	537	5	AY389971	AY389971 Xenopus l
C 394	22	91.7	479	6	AX435985	AX435985 Sequence	C 467	22	91.7	537	1	AY476797	AY476797 Unculture
C 395	22	91.7	479	8	AK062820	AK062820 Oryza sat	C 468	22	91.7	540	10	RATNUCBA4	M37038 Rat nucleol
C 396	22	91.7	480	3	AF502083	AF502083 Aphiz gos	469	22	91.7	540	10	RATNUCBA4	M37038 Rat nucleol
C 397	22	91.7	480	6	AL37333	AL37333 variable re	470	22	91.7	546	6	CQ719372	CQ719372 Sequence
C 398	22	91.7	482	11	BV070452	BV070452 S208P6807	471	22	91.7	546	6	CQ719372	CQ719372 Sequence
C 399	22	91.7	483	1	UAL240922	UAL240922 unculture	C 472	22	91.7	546	6	BD147974	BD147974 Primer fo
C 400	22	91.7	484	6	AR422708	AR422708 Sequence	C 473	22	91.7	546	6	BD147974	BD147974 Primer fo
C 401	22	91.7	484	6	AX198093	AX198093 Sequence	C 474	22	91.7	546	6	BD147974	BD147974 Primer fo
C 402	22	91.7	484	6	AX983402	AX983402 Sequence	C 475	22	91.7	546	6	BD147974	BD147974 Primer fo
C 403	22	91.7	484	6	BD118261	BD118261 EST and e	C 476	22	91.7	546	6	BD147974	BD147974 Primer fo
C 404	22	91.7	485	5	AF4659985	AF4659985 Oncorhync	C 477	22	91.7	552	3	AY196860	AY196860 Otiiorhync
C 405	22	91.7	486	6	CQ406779	CQ406779 Sequence	C 478	22	91.7	552	3	AY196862	AY196862 Otiiorhync
C 406	22	91.7	486	6	AY476782	AY476782 Unculture	C 479	22	91.7	552	3	AY196863	AY196863 Otiiorhync
C 407	22	91.7	486	6	AX197702	AX197702 Sequence	C 480	22	91.7	552	3	AY196865	AY196865 Otiiorhync
C 408	22	91.7	486	6	AX779040	AX779040 Sequence	C 481	22	91.7	552	3	AY196867	AY196867 Otiiorhync
C 409	22	91.7	488	11	GS3016	GS3016 SHGC-140787	C 482	22	91.7	552	3	AY196868	AY196868 Otiiorhync
C 410	22	91.7	490	6	AX306056	AX306056 Sequence	C 483	22	91.7	552	3	AY196868	AY196868 Otiiorhync
C 411	22	91.7	491	1	AF522934	AF522934 Unculture	C 484	22	91.7	552	3	AY196869	AY196869 Otiiorhync
C 412	22	91.7	491	4	AY434764	AY434764 Felis cat	C 485	22	91.7	552	3	AY196870	AY196870 Otiiorhync
C 413	22	91.7	492	1	AY193270	AY193270 Unculture	C 486	22	91.7	552	3	AY196871	AY196871 Otiiorhync
C 414	22	91.7	492	6	CQ056177	CQ056177 Sequence	C 487	22	91.7	552	3	AY196872	AY196872 Otiiorhync
C 415	22	91.7	492	6	CQ075435	CQ075435 Sequence	C 488	22	91.7	552	3	AY196873	AY196873 Otiiorhync
C 416	22	91.7	492	6	CQ106405	CQ106405 Sequence	C 489	22	91.7	552	3	AY196874	AY196874 Otiiorhync
C 417	22	91.7	492	6	CQ145076	CQ145076 Sequence	C 490	22	91.7	552	3	AY196875	AY196875 Strophoso
C 418	22	91.7	492	6	CQ180516	CQ180516 Sequence	C 491	22	91.7	552	3	AY196876	AY196876 Otiiorhync
C 419	22	91.7	492	6	CQ204898	CQ204898 Sequence	C 492	22	91.7	552	3	AY280621	AY280621 Otiiorhync
C 420	22	91.7	492	6	CQ228255	CQ228255 Sequence	C 493	22	91.7	554	1	AY476740	AY476740 Unculture
C 421	22	91.7	492	6	CQ266429	CQ266429 Sequence	C 494	22	91.7	555	6	AR134677	AR134677 Sequence
C 422	22	91.7	492	6	CQ303402	CQ303402 Sequence	495	22	91.7	555	6	AR152038	AR152038 Sequence
C 423	22	91.7	492	6	CQ340697	CQ340697 Sequence	C 496	22	91.7	556	6	CQ102996	CQ102996 Sequence
C 424	22	91.7	495	1	AF392763	AF392763 Unculture	497	22	91.7	556	6	CQ141862	CQ141862 Sequence
C 425	22	91.7	495	6	AR392765	AR392765 Unculture	C 498	22	91.7	556	6	CQ263172	CQ263172 Sequence
C 426	22	91.7	495	6	AR395323	AR395323 Sequence	499	22	91.7	556	6	CQ300205	CQ300205 Sequence
C 427	22	91.7	497	1	AY49426	AY49426 Unculture	C 500	22	91.7	556	6	CQ300205	CQ300205 Sequence
C 428	22	91.7	497	1	AF547404	AF547404 Unculture	501	22	91.7	556	6	CQ337454	CQ337454 Sequence
C 429	22	91.7	498	1	AF547404	AF547404 Unculture	C 502	22	91.7	558	1	AF245621	AF245621 Unculture
C 430	22	91.7	499	1	AF392780	AF392780 Unculture	C 503	22	91.7	559	8	AK061707	AK061707 Oryza sat
C 431	22	91.7	499	1	AF392780	AF392780 Unculture	504	22	91.7	560	11	BV048030	BV048030 S209P6512
C 432	22	91.7	500	1	AB058672	AB058672 Unculture	C 505	22	91.7	565	6	CQ527229	CQ527229 Sequence
C 433	22	91.7	502	6	CQ100207	CQ100207 Sequence	506	22	91.7	566	6	AX871907	AX871907 Sequence
C 434	22	91.7	502	6	CQ139195	CQ139195 Sequence	C 507	22	91.7	566	6	BD151969	BD151969 Primer fo
C 435	22	91.7	502	6	CQ222584	CQ222584 Sequence	508	22	91.7	567	6	CQ408937	CQ408937 Sequence
C 436	22	91.7	502	6	CQ260538	CQ260538 Sequence	C 509	22	91.7	567	6	CQ581442	CQ581442 Sequence
C 437	22	91.7	502	6	CQ334670	CQ334670 Sequence	510	22	91.7	568	6	CQ100881	CQ100881 Sequence
C 438	22	91.7	503	14	AY584590	AY584590 Nootka lu	C 511	22	91.7	568	6	CQ139876	CQ139876 Sequence
C 439	22	91.7	504	6	AX698138	AX698138 Sequence	C 512	22	91.7	568	6	CQ223230	CQ223230 Sequence
C 440	22	91.7	504	9	HUMHPP16B	M15780 Human DNA/e	C 513	22	91.7	568	6	CQ261198	CQ261198 Sequence
C 441	22	91.7	507	6	AX433363	AX433363 Sequence	C 514	22	91.7	568	6	CQ335346	CQ335346 Sequence
C 442	22	91.7	509	1	AY337356	AY337356 Paracoccu	C 515	22	91.7	569	6	AX509404	AX509404 Sequence

516	22	91.7	569	6	AX510030	AX510030 Sequence	c 589	22	91.7	650	8	AF374536	AF374536 Thalassio
517	22	91.7	572	11	BV016260	BV016260 RPAMSEQ0	c 590	22	91.7	650	9	AF399524	AF399524 Homo sapi
518	22	91.7	573	6	CQ408757	CQ408757 Sequence	c 591	22	91.7	650	11	BV049357	BV049357 S212P6018
519	22	91.7	575	8	AV538328	AV538328 Ipomoea t	c 592	22	91.7	651	9	HS333582	HS333582 Homo sapi
520	22	91.7	575	11	BV072514	BV072514 S212P6020	c 593	22	91.7	651	11	BV032970	BV032970 S212P6616
521	22	91.7	577	1	EC0HPR1	XO4128 E. coli pla	c 594	22	91.7	652	8	AF159798	AF159798 Arabidops
522	22	91.7	577	6	CQ072759	CQ072759 Sequence	c 595	22	91.7	653	11	BV042164	BV042164 S212P6030
523	22	91.7	579	6	CQ103566	CQ103566 Sequence	c 596	22	91.7	654	1	AY159800	AY159800 Paracoccu
524	22	91.7	579	6	CQ142400	CQ142400 Sequence	c 597	22	91.7	654	8	AF374537	AF374537 Thalassio
525	22	91.7	579	6	CQ300747	CQ300747 Sequence	c 598	22	91.7	654	8	AF374538	AF374538 Thalassio
526	22	91.7	581	6	CQ409345	CQ409345 Sequence	c 599	22	91.7	654	8	AF374539	AF374539 Thalassio
527	22	91.7	581	11	G81112	G81112 S209P6227RF	c 600	22	91.7	654	11	G55859	G55859 SHGC-101262
528	22	91.7	582	11	BV074683	BV074683 S209P6108	c 601	22	91.7	655	11	BV073566	BV073566 S212P6032
529	22	91.7	584	6	CQ523381	CQ523381 Sequence	c 602	22	91.7	655	11	BV073566	AX869324 Sequence
530	22	91.7	586	1	AY476761	AY476761 Unculture	c 603	22	91.7	656	6	BD149386	BD149386 Primer fo
531	22	91.7	586	9	HSU08107	U08107 Human N-net	c 604	22	91.7	657	8	AF121142	AF121142 Gracilari
532	22	91.7	587	8	PTR286129	AJ286129 Populus t	c 605	22	91.7	657	9	HS3331089	AJ3331089 Homo sapi
533	22	91.7	588	1	AY476794	AY476794 Unculture	c 606	22	91.7	662	1	AY332205	AY332205 Alpha pro
534	22	91.7	588	11	G98409	G98409 S210P6176RE	c 607	22	91.7	662	10	AY170535	AY170535 Mus muscu
535	22	91.7	591	6	AX400700	AX400700 Sequence	c 608	22	91.7	663	1	AY332135	AY332135 Alpha pro
536	22	91.7	592	1	AF245631	AF245631 Unculture	c 609	22	91.7	663	11	PM12C5B	AL684343 Penicilli
537	22	91.7	592	6	CQ524158	CQ524158 Sequence	c 610	22	91.7	664	1	AY332194	AY332194 Alpha pro
538	22	91.7	594	11	BV161552	BV161552 RPAMSEQ0	c 611	22	91.7	665	6	CQ432559	CQ432559 Sequence
539	22	91.7	596	6	CQ669824	CQ669824 Sequence	c 612	22	91.7	665	6	BD267075	BD267075 Composi
540	22	91.7	597	6	AX415608	AX415608 Sequence	c 613	22	91.7	667	9	AY448887	AY448887 Cercopith
541	22	91.7	597	6	AX415608	AX415608 Sequence	c 614	22	91.7	670	6	BD220648	BD220648 Human gen
542	22	91.7	597	6	AX972442	AX972442 Sequence	c 615	22	91.7	671	3	AF082189	AF082189 Haemopis
543	22	91.7	597	6	BD111161	BD111161 EST and e	c 616	22	91.7	673	8	AF348007	AF348007 Cooktowni
544	22	91.7	601	8	AY052153	AY052153 Triticum	c 617	22	91.7	674	1	AY332120	AY332120 Mesorhizo
545	22	91.7	603	6	AX123500	AX123500 Sequence	c 618	22	91.7	674	8	AF348035	AF348035 Habenaria
546	22	91.7	603	6	BD165617	BD165617 Novel pol	c 619	22	91.7	675	8	BT014807	BT014807 Arabidops
547	22	91.7	604	11	G89455	G89455 S209P6336FH	c 620	22	91.7	678	11	BV061031	BV061031 S212P6310
548	22	91.7	605	8	ATHSEQA	M80566 A.thaliana	c 621	22	91.7	680	4	AF513746	AF513746 Barbastei
549	22	91.7	605	11	BV002006	BV002006 S208P6553	c 622	22	91.7	680	4	AF513750	AF513750 Barbastei
550	22	91.7	605	11	G87144	G87144 S210P6133RE	c 623	22	91.7	680	4	AF513753	AF513753 Barbastei
551	22	91.7	606	6	CQ751969	CQ751969 Sequence	c 624	22	91.7	680	4	AF348019	AF348019 Cynorkis
552	22	91.7	609	1	PR429231	AJ299231 Paracoccu	c 625	22	91.7	681	11	BV042227	BV042227 S212P6005
553	22	91.7	611	11	AV525625	AV525625 Unculture	c 626	22	91.7	684	11	BV030942	BV030942 S212P6045
554	22	91.7	613	11	BV057922	BV057922 S212P6039	c 627	22	91.7	685	1	AY332157	AY332157 Paracoccu
555	22	91.7	614	6	BD229343	BD229343 Human gen	c 628	22	91.7	685	14	BYM224636	AJ224636 Barley ye
556	22	91.7	614	11	BV073298	BV073298 S212P6037	c 629	22	91.7	687	1	AF489286	AF489286 Bacterium
557	22	91.7	620	9	HS3329863	AJ329863 Homo sapi	c 630	22	91.7	687	1	AY193268	AY193268 Unculture
558	22	91.7	621	6	CQ595767	CQ595767 Sequence	c 631	22	91.7	687	11	BV022711	BV022711 S212P6808
559	22	91.7	622	11	G93830	G93830 S209P6221FH	c 632	22	91.7	688	1	AY332127	AY332127 Mesorhizo
560	22	91.7	623	11	BV049521	BV049521 S212P6010	c 633	22	91.7	691	1	AY515429	AY515429 Unculture
561	22	91.7	623	11	G86206	G86206 S208P6399RE	c 634	22	91.7	691	11	BV074246	BV074246 S212P6067
562	22	91.7	625	6	CQ780120	CQ780120 Sequence	c 635	22	91.7	692	1	AY193271	AY193271 Unculture
563	22	91.7	625	6	BD124829	BD124829 Primer fo	c 636	22	91.7	693	8	AF052865	AF052865 Peperomia
564	22	91.7	626	6	CQ396650	CQ396650 Sequence	c 637	22	91.7	695	11	AY332145	AY332145 Alpha pro
565	22	91.7	627	5	CQ402961	CQ402961 Sequence	c 638	22	91.7	695	3	AY060855	AY060855 Drosophil
566	22	91.7	627	5	AV463243	AV463243 Meleagris	c 639	22	91.7	696	8	CNS01DNL	AL117049 Batrytis
567	22	91.7	627	5	AV463243	AV463243 Meleagris	c 640	22	91.7	696	1	AY032949	AY032949 Paenibaci
568	22	91.7	629	1	AV337355	AV337355 Paracoccu	c 641	22	91.7	698	1	AY032949	AY032949 Paenibaci
569	22	91.7	629	3	PU06382	AJ006382 Platythr	c 642	22	91.7	698	1	AY193225	AY193225 Unculture
570	22	91.7	631	1	URH548915	AJ548915 unculture	c 643	22	91.7	698	1	AY193232	AY193232 Unculture
571	22	91.7	632	11	BV102747	BV102747 MARC 2135	c 644	22	91.7	698	1	AY193233	AY193233 Unculture
572	22	91.7	634	8	AF049296	AF049296 Arabidops	c 645	22	91.7	699	3	AF169744	AF169744 Paratanai
573	22	91.7	636	5	AX232227	AX232227 Squallus a	c 646	22	91.7	699	1	AY193227	AY193227 Unculture
574	22	91.7	637	1	AF017118	AF017118 Chromatiu	c 647	22	91.7	699	6	AR347826	AR347826 Sequence
575	22	91.7	640	6	AX371014	AX371014 Sequence	c 648	22	91.7	700	6	AX182322	AX182322 Sequence
576	22	91.7	641	11	BV022075	BV022075 S212P6116	c 649	22	91.7	700	6	AX182854	AX182854 Sequence
577	22	91.7	641	11	BV035185	BV035185 S212P6243	c 650	22	91.7	700	6	AX182855	AX182855 Sequence
578	22	91.7	642	6	AR134683	AR134683 Sequence	c 651	22	91.7	702	1	AY193226	AY193226 Unculture
579	22	91.7	642	6	AR152044	AR152044 Sequence	c 652	22	91.7	703	6	AX414756	AX414756 Sequence
580	22	91.7	642	6	AR242269	AR242269 Sequence	c 653	22	91.7	703	11	BV013173	BV013173 S212P6017
581	22	91.7	642	6	AX481403	AX481403 Sequence	c 654	22	91.7	705	1	AY193266	AY193266 Unculture
582	22	91.7	644	11	G97143	G97143 S208P6440FF	c 655	22	91.7	705	1	AY592763	AY592763 Unculture
583	22	91.7	644	1	AY102910	AY102910 Unculture	c 656	22	91.7	705	8	AF693041	AF693041 Saccharom
584	22	91.7	647	5	AX463242	AX463242 Gallus ga	c 657	22	91.7	705	11	BV013345	BV013345 S212P6019
585	22	91.7	648	8	AB105237	AB105237 Rhododend	c 658	22	91.7	708	6	BD214758	BD214758 Novel hum
586	22	91.7	649	1	AY332139	AY332139 Alpha pro	c 659	22	91.7	708	14	AB090801	AB090801 Human cox
587	22	91.7	650	8	AF374534	AF374534 Thalassio	c 660	22	91.7	709	3	BC075148	BC075148 Macrobac
588	22	91.7	650	8	AF374535	AF374535 Thalassio	c 661	22	91.7	709	5	BC075140	BC075140 Xenopus 1

c 662	22	91.7	709	5	MCU77596	U77596 Micrurus co	735	22	91.7	785	6	CQ425911	CQ425911 Sequence
c 663	22	91.7	710	11	BV046127	BV046127 S212P6016	c 736	22	91.7	785	8	MZEPYRDE2	D14457 Zea mays mR
c 664	22	91.7	712	1	AY102890	AY102890 Unculture	c 737	22	91.7	785	8	ZMPDC3	Z21722 Z.mays mRNa
c 665	22	91.7	712	1	AY193280	AY193280 Unculture	c 738	22	91.7	785	8	AK058546	AK058546 Oryza sat
c 666	22	91.7	712	1	AY193280	AY193280 Unculture	c 739	22	91.7	785	8	AK058546	AK058546 Oryza sat
c 667	22	91.7	714	6	CQ183480	CQ183480 Unculture	c 740	22	91.7	786	14	AF548619	AF548619 Newcastl
c 668	22	91.7	717	1	AY437623	AY437623 Paracoccu	c 741	22	91.7	789	11	BV050032	BV050032 S212P6600
c 669	22	91.7	717	9	HA332201	HA332201 Homo sapi	c 742	22	91.7	791	1	AY580458	AY580458 Unculture
c 670	22	91.7	718	6	AK065481	AK065481 Sequence	c 743	22	91.7	791	1	AY580441	AY580441 Unculture
c 671	22	91.7	721	5	CR388739	CR388739 Gallus ga	c 744	22	91.7	793	1	AY580445	AY580445 Unculture
c 672	22	91.7	722	5	CR388739	CR388739 Gallus ga	c 745	22	91.7	793	1	AY580468	AY580468 Unculture
c 673	22	91.7	724	11	PM4C1G	PM4C1G Penicilliu	c 746	22	91.7	800	14	AY062234	AY062234 Unculture
c 674	22	91.7	725	11	BV061734	BV061734 S212P6080	c 747	22	91.7	804	14	AY370771	AY370771 Newcastl
c 675	22	91.7	726	6	BD229034	BD229034 Genes and	c 748	22	91.7	810	6	CQ417232	CQ417232 Sequence
c 676	22	91.7	726	6	CQ584082	CQ584082 Sequence	c 749	22	91.7	814	6	AY580435	AY580435 Unculture
c 677	22	91.7	726	6	AR244261	AR244261 Sequence	c 750	22	91.7	814	6	CQ413799	CQ413799 Sequence
c 678	22	91.7	731	6	AR447616	AR447616 Sequence	c 751	22	91.7	814	6	AK110744	AK110744 Oryza sat
c 679	22	91.7	731	6	AX756226	AX756226 Sequence	c 752	22	91.7	816	9	HS337172	HS337172 Homo sapi
c 680	22	91.7	732	1	AY102903	AY102903 Unculture	c 753	22	91.7	817	3	AF382058	AF382058 Pseudohal
c 681	22	91.7	736	1	UBZ88581	UBZ88581 Unidentifie	c 754	22	91.7	817	3	AF382058	AF382058 Pseudohal
c 682	22	91.7	737	1	UBZ88582	UBZ88582 Unidentifie	c 755	22	91.7	818	3	S79577	S79577 K99 cluster
c 683	22	91.7	738	1	PLY15324	PLY15324 Paracoccu	c 756	22	91.7	818	3	AF382058	AF382058 Pseudohal
c 684	22	91.7	738	6	AX000501	AX000501 Sequence	c 757	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 685	22	91.7	738	6	BD073335	BD073335 Gene for	c 758	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 686	22	91.7	740	8	ATH553429	ATH553429 Arabidops	c 759	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 687	22	91.7	741	1	AY327152	AY327152 Unculture	c 760	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 688	22	91.7	742	1	AY580450	AY580450 Unculture	c 761	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 689	22	91.7	742	6	AX869801	AX869801 Sequence	c 762	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 690	22	91.7	742	6	BD149863	BD149863 Primer fo	c 763	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 691	22	91.7	743	6	BD229033	BD229033 Genes and	c 764	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 692	22	91.7	743	6	CQ780016	CQ780016 Sequence	c 765	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 693	22	91.7	743	6	CQ782205	CQ782205 Sequence	c 766	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 694	22	91.7	743	6	AR244260	AR244260 Sequence	c 767	22	91.7	819	3	AF382058	AF382058 Pseudohal
c 695	22	91.7	743	6	BD124725	BD124725 Primer fo	c 768	22	91.7	820	1	AY360492	AY360492 Unculture
c 696	22	91.7	743	6	BD126914	BD126914 Primer fo	c 769	22	91.7	824	1	AY176598	AY176598 Unculture
c 697	22	91.7	745	1	AY268232	AY268232 Unculture	c 770	22	91.7	824	1	AY176598	AY176598 Unculture
c 698	22	91.7	745	6	CQ782163	CQ782163 Sequence	c 771	22	91.7	827	11	BV014474	BV014474 S212P6016
c 699	22	91.7	746	6	AX870340	AX870340 Sequence	c 772	22	91.7	827	11	BV014474	BV014474 S212P6016
c 700	22	91.7	746	6	BD126872	BD126872 Primer fo	c 773	22	91.7	830	1	AB119499	AB119499 Halomonas
c 701	22	91.7	746	6	BD150402	BD150402 Primer fo	c 774	22	91.7	830	1	AB119499	AB119499 Halomonas
c 702	22	91.7	749	1	MAU512210	MAU512210 Marine al	c 775	22	91.7	833	3	MLU40203	MLU40203 Monocellis
c 703	22	91.7	751	1	AY268244	AY268244 Unculture	c 776	22	91.7	836	3	MLU40203	MLU40203 Monocellis
c 704	22	91.7	751	5	CR390798	CR390798 Gallus ga	c 777	22	91.7	839	1	AY580469	AY580469 Unculture
c 705	22	91.7	752	6	AR319052	AR319052 Sequence	c 778	22	91.7	839	1	AY580469	AY580469 Unculture
c 706	22	91.7	753	6	AY319052	AY319052 Sequence	c 779	22	91.7	841	1	AB094833	AB094833 Unculture
c 707	22	91.7	753	9	AY53037S2	AY53037S2 Homo sapi	c 780	22	91.7	841	1	AB094833	AB094833 Unculture
c 708	22	91.7	754	1	AY580447	AY580447 Unculture	c 781	22	91.7	841	1	AB094833	AB094833 Unculture
c 709	22	91.7	754	6	AR155604	AR155604 Sequence	c 782	22	91.7	841	1	AB094833	AB094833 Unculture
c 710	22	91.7	754	6	E66129	E66129 Genome DNA	c 783	22	91.7	841	1	AB094833	AB094833 Unculture
c 711	22	91.7	754	8	AF276805	AF276805 Silene la	c 784	22	91.7	841	1	AB094833	AB094833 Unculture
c 712	22	91.7	755	6	AR237210	AR237210 Metastati	c 785	22	91.7	841	1	AB094833	AB094833 Unculture
c 713	22	91.7	755	6	AR240775	AR240775 Sequence	c 786	22	91.7	841	1	AB094833	AB094833 Unculture
c 714	22	91.7	755	11	BV034791	BV034791 S212P6936	c 787	22	91.7	841	1	AB094833	AB094833 Unculture
c 715	22	91.7	755	11	AY177715	AY177715 Phyllobac	c 788	22	91.7	841	1	AB094833	AB094833 Unculture
c 716	22	91.7	762	1	AY193259	AY193259 Unculture	c 789	22	91.7	841	1	AB094833	AB094833 Unculture
c 717	22	91.7	762	1	AY193259	AY193259 Unculture	c 790	22	91.7	841	1	AB094833	AB094833 Unculture
c 718	22	91.7	765	9	HS327772	HS327772 Homo sapi	c 791	22	91.7	841	1	AB094833	AB094833 Unculture
c 719	22	91.7	766	1	AF327729	AF327729 Aeromonas	c 792	22	91.7	841	1	AB094833	AB094833 Unculture
c 720	22	91.7	766	1	AY193264	AY193264 Unculture	c 793	22	91.7	841	1	AB094833	AB094833 Unculture
c 721	22	91.7	767	11	BV039515	BV039515 S212P6454	c 794	22	91.7	841	1	AB094833	AB094833 Unculture
c 722	22	91.7	768	6	AR385113	AR385113 Sequence	c 795	22	91.7	841	1	AB094833	AB094833 Unculture
c 723	22	91.7	770	1	LEPSSR	LEPSSR Leptospira	c 796	22	91.7	841	1	AB094833	AB094833 Unculture
c 724	22	91.7	772	5	AY662689	AY662689 Gallus ga	c 797	22	91.7	841	1	AB094833	AB094833 Unculture
c 725	22	91.7	774	5	BRU16310	BRU16310 Brachydanio	c 798	22	91.7	841	1	AB094833	AB094833 Unculture
c 726	22	91.7	776	1	AB094948	AB094948 Unculture	c 799	22	91.7	841	1	AB094833	AB094833 Unculture
c 727	22	91.7	777	1	AY368462	AY368462 Unculture	c 800	22	91.7	841	1	AB094833	AB094833 Unculture
c 728	22	91.7	780	8	CNS01CFQ	CNS01CFQ Botrytis	c 801	22	91.7	841	1	AB094833	AB094833 Unculture
c 729	22	91.7	780	11	BV034886	BV034886 S212P6047	c 802	22	91.7	841	1	AB094833	AB094833 Unculture
c 730	22	91.7	781	10	MUSRSDI2B	MUSRSDI2B mouse repet	c 803	22	91.7	841	1	AB094833	AB094833 Unculture
c 731	22	91.7	782	11	BV075315	BV075315 S212P6052	c 804	22	91.7	841	1	AB094833	AB094833 Unculture
c 732	22	91.7	783	1	AJ630175	AJ630175 Paracoccu	c 805	22	91.7	841	1	AB094833	AB094833 Unculture
c 733	22	91.7	783	11	AY580470	AY580470 Unculture	c 806	22	91.7	841	1	AB094833	AB094833 Unculture
c 734	22	91.7	784	1	AY580470	AY580470 Unculture	c 807	22	91.7	841	1	AB094833	AB094833 Unculture

C 808	22	91.7	873	8	BT012980	Lycopersi	881	22	91.7	921	11	CNS061K6	AL400396 T3 end of
C 809	22	91.7	876	5	QULTR0FO1	M73702 Coturnix co	882	22	91.7	922	11	CNS061J4	AL400358 T7 end of
C 810	22	91.7	877	5	URH548908	AJ548908 unculture	883	22	91.7	923	8	PCU04135	U04135 Pleurotus c
C 811	22	91.7	877	5	BX935979	BX935979 Gallus ga	884	22	91.7	923	8	PCU04146	U04146 Pleurotus c
C 812	22	91.7	879	6	CQ412653	CQ412653 Sequence	885	22	91.7	923	8	AF261534	AF261534 Podocsyph
C 813	22	91.7	879	8	AF042629	AF042629 Campanell	886	22	91.7	924	6	AR484642	AR484642 Sequence
C 814	22	91.7	880	3	AF336288	AF336288 Galleria	887	22	91.7	924	6	AX143399	AX143399 Sequence
C 815	22	91.7	880	8	AZ19584	AZ19584 Gymnopili	c 888	22	91.7	924	8	TSP566653	AJ566653 Triticum
C 816	22	91.7	882	3	POU041734	POU041734 Ixodes sc	c 889	22	91.7	924	8	TSP566654	AJ566654 Triticum
C 817	22	91.7	886	8	POU04140	U04140 Pleurotus o	c 890	22	91.7	924	8	TSP566696	AJ566696 Triticum
C 818	22	91.7	886	8	POU04160	U04160 Pleurotus o	c 891	22	91.7	925	8	TAE566643	AJ566643 Triticum
C 819	22	91.7	886	8	POU04147	U04147 Pleurotus o	c 892	22	91.7	925	8	TAE566647	AJ566647 Triticum
C 820	22	91.7	887	3	BT015231	BT015231 Drosophil	c 893	22	91.7	925	8	TAE566651	AJ566651 Triticum
C 821	22	91.7	887	8	POU04143	U04143 Pleurotus o	c 894	22	91.7	925	8	TDS566655	AJ566655 Triticum
C 822	22	91.7	888	8	POU04144	U04144 Pleurotus o	c 895	22	91.7	925	8	TSP566642	AJ566642 Triticum
C 823	22	91.7	888	8	PFU04141	U04141 Pleurotus p	c 896	22	91.7	925	8	TSP566644	AJ566644 Triticum
C 824	22	91.7	888	8	PFU04145	U04145 Pleurotus p	c 897	22	91.7	925	8	TSP566645	AJ566645 Triticum
C 825	22	91.7	888	8	PFU04151	U04151 Pleurotus p	c 898	22	91.7	925	8	TSP566646	AJ566646 Triticum
C 826	22	91.7	888	8	PFU04157	U04157 Pleurotus p	c 899	22	91.7	925	8	TSP566648	AJ566648 Triticum
C 827	22	91.7	888	14	AY372673	AY372673 Newcastle	c 900	22	91.7	925	8	TSP566649	AJ566649 Triticum
C 828	22	91.7	889	8	PFU04152	U04152 Pleurotus p	c 901	22	91.7	925	8	TSP566650	AJ566650 Triticum
C 829	22	91.7	889	8	PFU04153	U04153 Pleurotus p	c 902	22	91.7	925	8	TSP566652	AJ566652 Triticum
C 830	22	91.7	890	1	AY277267	AY277267 Sulfitoba	c 903	22	91.7	925	8	TSP566698	AJ566698 Triticum
C 831	22	91.7	892	1	AY568815	AY568815 Unculture	c 904	22	91.7	926	5	BC071487	BC071487 Danio rer
C 832	22	91.7	892	8	POS406580	AJ406580 Pleurotus	c 905	22	91.7	926	8	FDU04139	U04139 Pleurotus d
C 833	22	91.7	894	6	AX616893	AX616893 Sequence	c 906	22	91.7	927	6	AR319436	AR319436 Sequence
C 834	22	91.7	894	6	AX952977	AX952977 Sequence	c 907	22	91.7	927	8	PCU04148	U04148 Pleurotus c
C 835	22	91.7	897	1	AY349532	AY349532 Unculture	c 908	22	91.7	927	8	PCU04149	U04149 Pleurotus c
C 836	22	91.7	897	8	D63888	D63888 Nicotiana a	c 909	22	91.7	927	8	TAE566691	AJ566691 Triticum
C 837	22	91.7	897	9	AF179877	AF179877 Homo sapi	c 910	22	91.7	927	8	TSP566665	AJ566665 Triticum
C 838	22	91.7	897	9	HS4322625	AJ322825 Homo sapi	c 911	22	91.7	927	8	TSP566666	AJ566666 Triticum
C 839	22	91.7	898	1	URH548896	AJ548896 unculture	c 912	22	91.7	927	8	TSP566680	AJ566680 Triticum
C 840	22	91.7	898	4	AY345584	AY345584 Canis fam	c 913	22	91.7	927	8	TSP566692	AJ566692 Triticum
C 841	22	91.7	900	6	CQ579012	CQ579012 Sequence	c 914	22	91.7	927	8	TSP566693	AJ566693 Triticum
C 842	22	91.7	901	8	AY085793	AY085793 Arabidops	c 915	22	91.7	927	8	TSP566694	AJ566694 Triticum
C 843	22	91.7	901	9	AF545815	AF545815 Homo sapi	c 916	22	91.7	928	8	FDU04156	U04156 Pleurotus d
C 844	22	91.7	902	8	PFU04154	U04154 Pleurotus e	c 917	22	91.7	928	8	FSU04150	U04150 Pleurotus s
C 845	22	91.7	902	8	PFU04159	U04159 Pleurotus p	c 918	22	91.7	928	8	TAE566660	AJ566660 Triticum
C 846	22	91.7	902	8	AY293191	AY293191 Laetiporu	c 919	22	91.7	928	8	TAE566661	AJ566661 Triticum
C 847	22	91.7	903	14	AY372671	AY372671 Newcastle	c 920	22	91.7	928	8	TAE566664	AJ566664 Triticum
C 848	22	91.7	903	14	AY372672	AY372672 Newcastle	c 921	22	91.7	928	8	TAE566673	AJ566673 Triticum
C 849	22	91.7	904	8	AF261365	AF261365 Gerronema	c 922	22	91.7	928	8	TAE566682	AJ566682 Triticum
C 850	22	91.7	905	8	AT056684	AF205684 Crepidotu	c 923	22	91.7	928	8	TAE566684	AJ566684 Triticum
C 851	22	91.7	906	8	AY293200	AY293200 Phlebia t	c 924	22	91.7	928	8	TAE566688	AJ566688 Triticum
C 852	22	91.7	907	8	PFU04155	U04155 Pleurotus e	c 925	22	91.7	928	8	TDS566656	AJ566656 Triticum
C 853	22	91.7	907	8	POU04142	U04142 Pleurotus o	c 926	22	91.7	928	8	TDS566657	AJ566657 Triticum
C 854	22	91.7	908	8	PFU04136	U04136 Pleurotus f	c 927	22	91.7	928	8	TSP566658	AJ566658 Triticum
C 855	22	91.7	908	8	PFU04137	U04137 Pleurotus f	c 928	22	91.7	928	8	TSP566659	AJ566659 Triticum
C 856	22	91.7	909	3	GIU94701	U94701 Giardia int	c 929	22	91.7	928	8	TSP566662	AJ566662 Triticum
C 857	22	91.7	909	6	AX380436	AX380436 Sequence	c 930	22	91.7	928	8	TSP566663	AJ566663 Triticum
C 858	22	91.7	910	8	AY458787	AY458787 Hordeum b	c 931	22	91.7	928	8	TSP566667	AJ566667 Triticum
C 859	22	91.7	910	8	AY458789	AY458789 Hordeum b	c 932	22	91.7	928	8	TSP566668	AJ566668 Triticum
C 860	22	91.7	910	8	AY458793	AY458793 Hordeum j	c 933	22	91.7	928	8	TSP566669	AJ566669 Triticum
C 861	22	91.7	911	8	MAU406535	AJ406535 Mycoacia	c 934	22	91.7	928	8	TSP566670	AJ566670 Triticum
C 862	22	91.7	911	8	AF261338	AF261338 Tetrapyrg	c 935	22	91.7	928	8	TSP566671	AJ566671 Triticum
C 863	22	91.7	912	8	PSP406568	AJ406568 Peniophor	c 936	22	91.7	928	8	TSP566672	AJ566672 Triticum
C 864	22	91.7	912	8	SHY406573	AJ406573 Scopuloid	c 937	22	91.7	928	8	TSP566674	AJ566674 Triticum
C 865	22	91.7	912	8	SHY406574	AJ406574 Scopuloid	c 938	22	91.7	928	8	TSP566675	AJ566675 Triticum
C 866	22	91.7	913	1	URH548912	AJ548912 unculture	c 939	22	91.7	928	8	TSP566676	AJ566676 Triticum
C 867	22	91.7	913	5	BX9331447	BX9331447 Gallus a	c 940	22	91.7	928	8	TSP566677	AJ566677 Triticum
C 868	22	91.7	914	8	AB1406527	AJ406527 Abortipor	c 941	22	91.7	928	8	TSP566678	AJ566678 Triticum
C 869	22	91.7	914	8	CJU406561	AJ406561 Campanell	c 942	22	91.7	928	8	TSP566679	AJ566679 Triticum
C 870	22	91.7	915	1	AY277266	AY277266 Sulfitoba	c 943	22	91.7	928	8	TSP566681	AJ566681 Triticum
C 871	22	91.7	915	6	AX653757	AX653757 Sequence	c 944	22	91.7	928	8	TSP566683	AJ566683 Triticum
C 872	22	91.7	915	8	AXH308485	AJ308485 Arabidops	c 945	22	91.7	928	8	TSP566685	AJ566685 Triticum
C 873	22	91.7	915	8	AY028793	AY028793 Hordeum m	c 946	22	91.7	928	8	TSP566686	AJ566686 Triticum
C 874	22	91.7	915	8	AY028795	AY028795 Hordeum m	c 947	22	91.7	928	8	TSP566687	AJ566687 Triticum
C 875	22	91.7	915	8	AY458783	AY458783 Hordeum m	c 948	22	91.7	928	8	TSP566689	AJ566689 Triticum
C 876	22	91.7	915	8	AY458784	AY458784 Hordeum m	c 949	22	91.7	928	8	TSP566690	AJ566690 Triticum
C 877	22	91.7	915	8	AY458790	AY458790 Hordeum b	c 950	22	91.7	928	8	TSP566695	AJ566695 Triticum
C 878	22	91.7	915	8	AY458796	AY458796 Hordeum s	c 951	22	91.7	929	1	AY568887	AY568887 Unculture
C 879	22	91.7	916	8	AF261432	AF261432 Pleurotus	c 952	22	91.7	929	8	AF042575	AF042575 Pleurotus
C 880	22	91.7	921	8	AF135178	AF135178 Pleurotus	c 953	22	91.7	929	8	AF360202	AF360202 Arabidops





```

Query Match:      91.67%      Indels:      0
DB:               6          Gaps:         0

US-10-030-194A-6 (1-6) x AR307357 (1-21)

QY  1  GlyTyr***ValGlu 5
DB  19 GGCTACACTGTGGAG 5

RESULT 4
LOCUS      BD070798          21 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Method to diagnose and treat pathological conditions resulting from
            deficient ion transport such as Pseudohypaldosteronism type-1.
ACCESSION  BD070798
VERSION    JP 2001514521-A/37.
KEYWORDS   unidentified
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 21)
AUTHORS    Lifton,R.P., Chang,S.S. and Rossier,B.C.
TITLE      Method to diagnose and treat pathological conditions resulting from
            deficient ion transport such as Pseudohypaldosteronism type-1
JOURNAL    Patent: JP 2001514521-A 37 11-SEP-2001;
COMMENT    YALE UNIVERSITY
            OS Unidentified
            PN JP 2001514521-A/37
            PD 11-SEP-2001
            PF 11-MAR-1998 JP 1998539716
            PR 11-MAR-1997 US 60/040171
            PI RICHARD P LIFTON,SUE S CHANG,BERNARD C ROSSIER PC
            C12Q1/68,C07K16/18,C12N15/12,C12N5/10,C07K14/47 CC Strandedness:
            Single;
            CC Topology: Linear;
            FH Key /desc = 'primer' Location/Qualifiers
            FT source 1..21 /organism='Unidentified'
            FT source 1..21 /organism='Unidentified'

ORIGIN
Alignment Scores:
Pred. No.:      196      Length:      21
Score:          22.00    Matches:      4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match:    91.67% Indels:      0
DB:             6      Gaps:      0

US-10-030-194A-6 (1-6) x BD070800 (1-21)

QY  1  GlyTyr***ValGlu 5
DB  19 GGCTACACTGTGGAG 5

RESULT 6
LOCUS      AX135256          28 bp      DNA      linear      PAT 29-MAY-2001
DEFINITION Sequence 64 from Patent WO0132928.
ACCESSION  AX135256
VERSION    AX135256.1 GI:14271605
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS    Farr,S.
TITLE      Methods of determining individual hypersensitivity to an agent
JOURNAL    Patent: WO 0132928-A 64 10-MAY-2001;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES   Location/Qualifiers
            source 1..28
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      265      Length:      28
Score:          22.00    Matches:      4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match:    91.67% Indels:      0
DB:             6      Gaps:      0

US-10-030-194A-6 (1-6) x AX135256 (1-28)

QY  1  GlyTyr***ValGlu 5
DB  19 GGCTACACTGTGGAG 5

```

```

Query Match:      91.67%      Indels:      0
DB:               6          Gaps:         0

US-10-030-194A-6 (1-6) x AR307357 (1-21)

QY  1  GlyTyr***ValGlu 5
DB  19 GGCTACACTGTGGAG 5

RESULT 4
LOCUS      BD070798          21 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Method to diagnose and treat pathological conditions resulting from
            deficient ion transport such as Pseudohypaldosteronism type-1.
ACCESSION  BD070798
VERSION    JP 2001514521-A/37.
KEYWORDS   unidentified
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 21)
AUTHORS    Lifton,R.P., Chang,S.S. and Rossier,B.C.
TITLE      Method to diagnose and treat pathological conditions resulting from
            deficient ion transport such as Pseudohypaldosteronism type-1
JOURNAL    Patent: JP 2001514521-A 37 11-SEP-2001;
COMMENT    YALE UNIVERSITY
            OS Unidentified
            PN JP 2001514521-A/37
            PD 11-SEP-2001
            PF 11-MAR-1998 JP 1998539716
            PR 11-MAR-1997 US 60/040171
            PI RICHARD P LIFTON,SUE S CHANG,BERNARD C ROSSIER PC
            C12Q1/68,C07K16/18,C12N15/12,C12N5/10,C07K14/47 CC Strandedness:
            Single;
            CC Topology: Linear;
            FH Key /desc = 'primer' Location/Qualifiers
            FT source 1..21 /organism='Unidentified'
            FT source 1..21 /organism='Unidentified'

ORIGIN
Alignment Scores:
Pred. No.:      196      Length:      21
Score:          22.00    Matches:      4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match:    91.67% Indels:      0
DB:             6      Gaps:      0

US-10-030-194A-6 (1-6) x BD070798 (1-21)

QY  1  GlyTyr***ValGlu 5
DB  19 GGCTACACTGTGGAG 5

RESULT 5
LOCUS      BD070800          21 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Method to diagnose and treat pathological conditions resulting from
            deficient ion transport such as Pseudohypaldosteronism type-1.
ACCESSION  BD070800
VERSION    BD070800.1 GI:22616403
KEYWORDS   JP 2001514521-A/39.
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 21)
AUTHORS    Lifton,R.P., Chang,S.S. and Rossier,B.C.

```

```

Db          10 GGATATAGCGTCGAA 24
RESULT 7
AX554130
LOCUS      AX554130          31 bp    DNA          linear    PAT 27-NOV-2002
DEFINITION Sequence 153 from Patent WO02074799.
ACCESSION  AX554130
VERSION     AX554130.1 GI:25898067
KEYWORDS   .
SOURCE      synthetic construct
ORGANISM    artificial sequences.
REFERENCE   1
AUTHORS     Freyssiinet,G., Rang,C. and Frutos,R.
TITLE       Pepsin-sensitive modified bacillus thuringiensis insecticidal toxin
JOURNAL     Patent: WO 02074799-A 153 26-SEP-2002;
            AVENTIS CROPS SCIENCE S.A. (FR)
FEATURES   Location/Qualifiers
            source
            1..31
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="oligonucleotide 112"

ORIGIN
Alignment Scores: 295          Length: 31
Pred. No.: 22.00          Matches: 4
Score: 80.00%             Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67%       Gaps: 0
DB: 6

US-10-030-194A-6 (1-6) x AX554130 (1-31)
QY          1 GlyTyr***ValGlu 5
Db          15 GGTTACTCGTGAG 29

RESULT 8
AX554131
LOCUS      AX554131          35 bp    DNA          linear    PAT 27-NOV-2002
DEFINITION Sequence 154 from Patent WO02074799.
ACCESSION  AX554131
VERSION     AX554131.1 GI:25898068
KEYWORDS   .
SOURCE      synthetic construct
ORGANISM    synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS     Freyssiinet,G., Rang,C. and Frutos,R.
TITLE       Pepsin-sensitive modified bacillus thuringiensis insecticidal toxin
JOURNAL     Patent: WO 02074799-A 154 26-SEP-2002;
            AVENTIS CROPS SCIENCE S.A. (FR)
FEATURES   Location/Qualifiers
            source
            1..35
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="oligonucleotide 113"

ORIGIN
Alignment Scores: 334          Length: 35
Pred. No.: 22.00          Matches: 4
Score: 80.00%             Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67%       Gaps: 0
DB: 6

US-10-030-194A-6 (1-6) x AX554131 (1-35)
QY          1 GlyTyr***ValGlu 5

```

```

Db          5 GGTACTCGTGAG 19
RESULT 9
AR075824/c
LOCUS      AR075824          48 bp    DNA          linear    PAT 30-AUG-2000
DEFINITION Sequence 7 from patent US 5958700.
ACCESSION  AR075824
VERSION     AR075824.1 GI:10002570
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 48)
AUTHORS     Nadeau,J.G., Pitner,J.Bruce., Linn,C.Preston. and Schram,J.L.
TITLE       Detection of nucleic acids by fluorescence quenching
JOURNAL     Patent: US 5958700-A 7 28-SEP-1999;
            Location/Qualifiers
            source
            1..48
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Alignment Scores: 465          Length: 48
Pred. No.: 22.00          Matches: 4
Score: 80.00%             Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67%       Gaps: 0
DB: 6

US-10-030-194A-6 (1-6) x AR075824 (1-48)
QY          1 GlyTyr***ValGlu 5
Db          44 GGTACTCAGTAGAG 30

RESULT 10
E30457/c
LOCUS      E30457          48 bp    DNA          linear    PAT 18-JUN-2001
DEFINITION Method for detecting target nucleic acid sequence and
            oligonucleotide.
ACCESSION  E30457
VERSION     E30457.1 GI:13025614
KEYWORDS   JP 1999056380-A/7.
SOURCE      unidentified
            unclassified.
ORGANISM    unclassified.
            James,G.N., J,B.P., C.P.R. and James,L.S.
REFERENCE   1 (bases 1 to 48)
AUTHORS     Method for detecting target nucleic acid sequence and
            Patent: JP 1999056380-A 7 02-MAR-1999;
            BECTON DICKINSON & CO
JOURNAL     OS Unidentified
            PN JP 1999056380-A/7
COMMENT     PD 02-MAR-1999
            PF 29-MAY-1998 JP 1998166141
            PR 30-MAY-1997 US 08/865.675
            PI JAMES G NADEAU, J BLUCE PITONA, C PRESTON RIN, JAMES L SHURAMU PC
            C12N15/09,C12Q1/68,G01N33/50,G01N33/566,C12N15/00 CC
            Strandedness: Single;
            CC Topology: Linear;
            FH Key
            FT source
            1..48
            /organism="Unidentified".
            Location/Qualifiers
            source
            1..48
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"

ORIGIN
Alignment Scores:

```

```

Pred. No.: 465 Length: 48
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x E30457 (1-48)

Qy 1 GlyTyr***ValGlu 5
Db 44 GGTACTCAGTAGAG 30

RESULT 11
LOCUS CQ544105 60 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 13740 from Patent WO0210449.
ACCESSION CQ544105
VERSION CQ544105.1 GI:41510369
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
TITLE Oligonucleotide library for detecting rna transcripts and splice
variants that populate a transcriptome
JOURNAL Patent: WO 0210449-A 13740 07-FEB-2002;
Compugen Inc. (US)
FEATURES
source Location/Qualifiers
1..60
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 587 Length: 60
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ544105 (1-60)

Qy 1 GlyTyr***ValGlu 5
Db 46 GGTATGCAGTTGAG 60

RESULT 12
LOCUS CQ552569 60 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 22204 from Patent WO0210449.
ACCESSION CQ552569
VERSION CQ552569.1 GI:41518996
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
TITLE Oligonucleotide library for detecting rna transcripts and splice
variants that populate a transcriptome
JOURNAL Patent: WO 0210449-A 22204 07-FEB-2002;
Compugen Inc. (US)
FEATURES
source Location/Qualifiers
1..60
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

```

ORIGIN
Alignment Scores:
Pred. No.: 587 Length: 60
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ552569 (1-60)

Qy 1 GlyTyr***ValGlu 5
Db 38 GGCTATTCAGTGGAG 24

RESULT 13
LOCUS CQ533924 65 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 3559 from Patent WO0210449.
ACCESSION CQ533924
VERSION CQ533924.1 GI:41500188
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
TITLE Oligonucleotide library for detecting rna transcripts and splice
variants that populate a transcriptome
JOURNAL Patent: WO 0210449-A 3559 07-FEB-2002;
Compugen Inc. (US)
FEATURES
source Location/Qualifiers
1..65
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"

ORIGIN
Alignment Scores:
Pred. No.: 638 Length: 65
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ533924 (1-65)

Qy 1 GlyTyr***ValGlu 5
Db 40 GGTACAGTGTGGAA 54

RESULT 14
LOCUS AR357220 75 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 3338 from patent US 6593114.
ACCESSION AR357220
VERSION AR357220.1 GI:33763304
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 75)
AUTHORS Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and
Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6593114-A 3338 15-JUL-2003;
Compugen Inc. (US)
FEATURES
source Location/Qualifiers
1..75
/organism="unknown"

```

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores: 740 Length: 75  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 6

US-10-030-194A-6 (1-6) x AR357220 (1-75)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 23 GGCTATACGCTAGAA 9

RESULT 15  
AR035489 AR035489 95 bp DNA linear PAT 29-SEP-1999  
LOCUS Sequence 61 from patent US 5871902.  
DEFINITION AR035489  
ACCESSION AR035489  
VERSION AR035489.1 GI:5952157  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 95)  
AUTHORS Weininger,S. and Weininger,A.M.  
TITLE Sequence-specific detection of nucleic acid hybrids using a DNA-binding molecule or assembly capable of discriminating perfect hybrids from non-perfect hybrids  
JOURNAL Patent: US 5871902-A 61.16-FEB-1999;  
FEATURES Location/Qualifiers  
source 1..95  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores: 948 Length: 95  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 6

US-10-030-194A-6 (1-6) x AR035489 (1-95)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 20 GGATATACAGTGGAA 34

RESULT 16  
CQ001247 CQ001247 100 bp DNA linear PAT 16-JAN-2004  
LOCUS Sequence 12709 from Patent EP1260592.  
DEFINITION CQ001247  
ACCESSION CQ001247  
VERSION CQ001247.1 GI:41007885  
KEYWORDS  
SOURCE Escherichia coli  
ORGANISM Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.

REFERENCE 1  
AUTHORS Donner,H., Drescher,B., Huber,A. and Weber,J.  
TITLE Biochip  
JOURNAL Patent: EP 1260592-A 12709 27-NOV-2002;  
MWG -Biotech AG (DE)  
FEATURES Location/Qualifiers  
source 1..100  
/organism="Escherichia coli"  
/mol\_type="unassigned DNA"

/db\_xref="taxon:562"  
/note="dacb b3183 U00096 3326604\_\_3328037"

ORIGIN

Alignment Scores: 1e+03 Length: 100  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 6

US-10-030-194A-6 (1-6) x CQ001247 (1-100)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 5 GGCTATGCAGTAGAA 19

RESULT 17  
CQ001248 CQ001248 100 bp DNA linear PAT 16-JAN-2004  
LOCUS Sequence 12710 from Patent EP1260592.  
DEFINITION CQ001248  
ACCESSION CQ001248  
VERSION CQ001248.1 GI:41007886  
KEYWORDS  
SOURCE Escherichia coli  
ORGANISM Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.

REFERENCE 1  
AUTHORS Donner,H., Drescher,B., Huber,A. and Weber,J.  
TITLE Biochip  
JOURNAL Patent: EP 1260592-A 12710 27-NOV-2002;  
MWG -Biotech AG (DE)  
FEATURES Location/Qualifiers  
source 1..100  
/organism="Escherichia coli"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:562"

/note="dacb b3183 U00096 3326604\_\_3328037"

ORIGIN

Alignment Scores: 1e+03 Length: 100  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 6

US-10-030-194A-6 (1-6) x CQ001248 (1-100)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 54 GGCTATGCAGTAGAA 68

RESULT 18  
CQ001249 CQ001249 100 bp DNA linear PAT 16-JAN-2004  
LOCUS Sequence 12711 from Patent EP1260592.  
DEFINITION CQ001249  
ACCESSION CQ001249  
VERSION CQ001249.1 GI:41007887  
KEYWORDS  
SOURCE Escherichia coli  
ORGANISM Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.

REFERENCE 1  
AUTHORS Donner,H., Drescher,B., Huber,A. and Weber,J.  
TITLE Biochip  
JOURNAL Patent: EP 1260592-A 12711 27-NOV-2002;  
MWG -Biotech AG (DE)  
FEATURES Location/Qualifiers

```
source
1..100
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
/note="dacb b3183 U00096 3326604__3328037"

ORIGIN
Alignment Scores:
Pred. No.: 1e+03 Length: 100
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.00% Indels: 0
DB: 91.67% Gaps: 0

US-10-030-194A-6 (1-6) x CQ001249 (1-100)

QY 1 GlyTyr***ValGlu 5
Db 77 GGCTATGCGTAGAA 91

RESULT 19
AR366373 AR366373 100 bp DNA linear PAT 12-SEP-2003
LOCUS Sequence 14 from patent US 6329170.
ACCESSION AR366373
VERSION AR366373.1 GI:34598799
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 100)
Holmes,E.H. and Sherwood,A.L.
Nucleic acids and proteins of a rat ganglioside GM1-specific
.alpha.1.fwdarw.2fucosyltransferase and uses thereof
JOURNAL Patent: US 6329170-A 14 11-DEC-2001;
FEATURES Location/Qualifiers
source
1..100
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1e+03 Length: 100
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.00% Indels: 0
DB: 91.67% Gaps: 0

US-10-030-194A-6 (1-6) x AR366373 (1-100)

QY 1 GlyTyr***ValGlu 5
Db 11 GGCTACACCGTGAA 25

RESULT 20
AR366376 AR366376 100 bp DNA linear PAT 12-SEP-2003
LOCUS Sequence 17 from patent US 6329170.
ACCESSION AR366376
VERSION AR366376.1 GI:34598802
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 100)
Holmes,E.H. and Sherwood,A.L.
Nucleic acids and proteins of a rat ganglioside GM1-specific
.alpha.1.fwdarw.2fucosyltransferase and uses thereof
JOURNAL Patent: US 6329170-A 17 11-DEC-2001;
FEATURES Location/Qualifiers
source
1..100
/organism="unknown"
/mol_type="genomic DNA"

source
1..100
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
/note="dacb b3183 U00096 3326604__3328037"

ORIGIN
Alignment Scores:
Pred. No.: 1e+03 Length: 100
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.00% Indels: 0
DB: 91.67% Gaps: 0

US-10-030-194A-6 (1-6) x AR366376 (1-100)

QY 1 GlyTyr***ValGlu 5
Db 11 GGCTACACCGTGAA 25

RESULT 21
AR435569 AR435569 100 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 14 from patent US 6656714.
ACCESSION AR435569
VERSION AR435569.1 GI:40198534
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 100)
Holmes,E.H. and Sherwood,A.L.
Nucleic acids and proteins of a rat ganglioside GM1-specific
.alpha.1.fwdarw.2 fucosyltransferase and uses thereof
JOURNAL Patent: US 6656714-A 14 02-DEC-2003;
FEATURES Location/Qualifiers
source
1..100
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1e+03 Length: 100
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 80.00% Indels: 0
DB: 91.67% Gaps: 0

US-10-030-194A-6 (1-6) x AR435569 (1-100)

QY 1 GlyTyr***ValGlu 5
Db 11 GGCTACACCGTGAA 25

RESULT 22
AR435572 AR435572 100 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 17 from patent US 6656714.
ACCESSION AR435572
VERSION AR435572.1 GI:40198537
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 100)
Holmes,E.H. and Sherwood,A.L.
Nucleic acids and proteins of a rat ganglioside GM1-specific
.alpha.1.fwdarw.2 fucosyltransferase and uses thereof
JOURNAL Patent: US 6656714-A 17 02-DEC-2003;
FEATURES Location/Qualifiers
source
1..100
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
```

US-10-030-194A-6 (1-6) x AR435572 (1-100)		/db_xref="taxon:9606"	
ORIGIN			
Alignment Scores:			
Pred. No.:	1e+03	Length:	100
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	6	Gaps:	0
US-10-030-194A-6 (1-6) x Q0713868 (1-127)			
QY		1 GlyTyr***ValGlu 5	
Db		25 GGCTATAGTGTAGAG 11	
RESULT 23			
AX088741	AX088741	108 bp	DNA
LOCUS	Sequence 67 from Patent WO0114416.		linear
DEFINITION	AX088741		PAT 17-MAR-2001
ACCESSION	AX088741		
VERSION	AX088741.1	GI:13397537	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1		
AUTHORS	Neeper,M.P., McClements,W.L., Jansen,K.U., Schultz,L.D., Chen,L.		
TITLE	and Wang,X.M.		
JOURNAL	Synthetic human papillomavirus genes		
FEATURES	Patent: WO 0114416-A 67 01-MAR-2001;		
source	Merck & Co., Inc. (US)		
	Location/Qualifiers		
	1..108		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Codon-Optimized HPV16 E2 fragment"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.08e+03	Length:	108
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	6	Gaps:	0
US-10-030-194A-6 (1-6) x AX088741 (1-108)			
QY		1 GlyTyr***ValGlu 5	
Db		21 GGCTACACGTGGAG 35	
RESULT 24			
CO713868/c	CO713868/c	127 bp	DNA
LOCUS	Sequence 58794 from Patent WO02070737.		linear
DEFINITION	AX088741		PAT 03-FEB-2004
ACCESSION	AX088741		
VERSION	AX088741.1	GI:42274725	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Liew,C.C., Marshall,W.E. and Zhang,H.		
TITLE	Compositions and methods relating to osteoarthritis		
JOURNAL	Patent: WO 02070737-A 58794 12-SEP-2002;		
FEATURES	Chondrogene Inc. (CA)		
source	Location/Qualifiers		
	1..127		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
US-10-030-194A-6 (1-6) x Q0672500 (1-129)			
ORIGIN			
Alignment Scores:			
Pred. No.:	1.3e+03	Length:	129
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	6	Gaps:	0
US-10-030-194A-6 (1-6) x Q0672500 (1-129)			
QY		1 GlyTyr***ValGlu 5	
Db		103 GGCTATGCGGTGAG 117	
RESULT 26			
CO113946/c	CO113946/c	136 bp	DNA
LOCUS	Sequence 22805 from Patent WO0157272.		linear
DEFINITION	AX088741		PAT 21-JAN-2004
ACCESSION	AX088741		
VERSION	AX088741.1	GI:41083816	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.		
TITLE	Human genome-derived single exon nucleic acid probes useful for		
JOURNAL	analysis of gene expression in human placenta		
FEATURES	Patent: WO 0157272-A 22805 09-AUG-2001;		
source	Aeomica, Inc. (US)		
	Location/Qualifiers		
	1..136		
	/db_xref="taxon:9606"		

```
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AP000431.3-EXPRESSED IN PLACENTA, SIGNAL =
0.61-NT HIT: M18095.1, EVALUE 4.70e+00-SWISSPROT HIT:
P11253, EVALUE 4.10e+00"

ORIGIN

Alignment Scores:
Pred. No.: 1.38e+03 Length: 136
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ113946 (1-136)

QY 1 GlyTyr***ValGlu 5
||| ||| ||| |||
Db 127 GGTATTCCGTGGAA 113

RESULT 27
CQ152829 136 bp DNA linear PAT 21-JAN-2004
LOCUS
DEFINITION Sequence 22851 from Patent WO0157276.
ACCESSION CQ152829
VERSION CQ152829.1 GI:41160179
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
AUTHORS Human genome-derived single exon nucleic acid probes useful for
TITLE analysis of gene expression in human bone marrow
JOURNAL Patent: WO 0157276-A 22851 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source
1. .136
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AP000431.3-EXPRESSED IN BONE MARROW, SIGNAL
= 0.52-NT HIT: M18095.1, EVALUE 4.70e+00-SWISSPROT HIT:
P11253, EVALUE 4.10e+00"

ORIGIN

Alignment Scores:
Pred. No.: 1.38e+03 Length: 136
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ152829 (1-136)

QY 1 GlyTyr***ValGlu 5
||| ||| ||| |||
Db 127 GGTATTCCGTGGAA 113

RESULT 28
CQ236146 136 bp DNA linear PAT 21-JAN-2004
LOCUS
DEFINITION Sequence 22985 from Patent WO0157273.
ACCESSION CQ236146
VERSION CQ236146.1 GI:41219424
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
AUTHORS HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
(03.10.00)<150> US 60/236,359<151> 27 September 2000
(27.09.00)<150> US 60/234,687<151> 21 September 2000
(21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Listing Engine
Patent: WO 0157273-A 22985 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source
1. .136
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AP000431.3-EXPRESSED IN ADULT LIVER, SIGNAL
= 0.53-NT HIT: M18095.1, EVALUE 4.70e+00-SWISSPROT HIT:
P11253, EVALUE 4.10e+00"

ORIGIN

Alignment Scores:
Pred. No.: 1.38e+03 Length: 136
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ236146 (1-136)

QY 1 GlyTyr***ValGlu 5
||| ||| ||| |||
Db 127 GGTATTCCGTGGAA 113

RESULT 29
CQ273724 136 bp DNA linear PAT 23-JAN-2004
LOCUS
DEFINITION Sequence 21985 from Patent WO0157277.
ACCESSION CQ273724
VERSION CQ273724.1 GI:41246328
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
AUTHORS Human genome-derived single exon nucleic acid probes useful for
TITLE analysis of gene expression in human fetal liver
JOURNAL Patent: WO 0157277-A 21985 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source
1. .136
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AP000431.3-EXPRESSED IN FETAL LIVER, SIGNAL
= 0.63-NT HIT: M18095.1, EVALUE 4.70e+00-SWISSPROT HIT:
P11253, EVALUE 4.10e+00"

ORIGIN

Alignment Scores:
Pred. No.: 1.38e+03 Length: 136
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0
```

```
US-10-030-194A-6 (1-6) x CQ273724 (1-136)

QY 1 GlyTyr***ValGlu 5
Db 127 GGTATTCCGTGGAA 113

RESULT 30
CQ348024/c 136 bp DNA linear PAT 23-JAN-2004
LOCUS
DEFINITION Sequence 22118 from Patent WO0157275.
ACCESSION CQ348024
VERSION CQ348024.1 GI:41297095
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human brain
Patent: WO 0157275-A 22118 09-AUG-2001;
Aeonica, Inc. (US)
FEATURES
source
1..136
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="WAP TO AP000431.3-EXPRESSED IN BRAIN. SIGNAL =
0.57-NT HIT: M18095.1, EVALUE 4.70e+00-SWISSPROT HIT:
P11253, EVALUE 4.10e+00"

ORIGIN
Alignment Scores:
Pred. No.: 1.38e+03 Length: 136
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ348024 (1-136)

QY 1 GlyTyr***ValGlu 5
Db 127 GGTATTCCGTGGAA 113

RESULT 31
CQ832356 152 bp DNA linear PAT 29-JUL-2004
LOCUS
DEFINITION Sequence 20 from Patent WO2004058972.
ACCESSION CQ832356
VERSION CQ832356.1 GI:50831980
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Thiesen,H.J. and Lorenz,P.
Human autoantigens and use thereof
Patent: WO 2004058972-A 20 15-JUL-2004;
Thiesen, Hans-Juergen (DE); Lorenz, Peter (DE)
FEATURES
source
1..152
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.55e+03 Length: 152
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ832356 (1-152)

QY 1 GlyTyr***ValGlu 5
Db 19 GGATATGCTGTAGAG 33

RESULT 32
F369389S54/c 161 bp DNA linear ROD 22-AUG-2002
LOCUS
DEFINITION Mus musculus strain 129P3/J lipoprotein receptor-related protein
(Lrp1) gene, exon 54.
ACCESSION AF369442
VERSION AF369442.1 GI:115825060
KEYWORDS
SEGMENT
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (Bases 1 to 161)
Smeijers,L., Willems,S., Lauwers,A., Thirty,E., van Leuven,F. and
Roebroek,A.J.
Functional expression of murine LRP1 requires correction of Lrp1
cDNA sequences
Biochim Biophys. Acta 1577 (1), 155-158 (2002)
MEDLINE 22145789
PUBMED 12151109
REFERENCE
2 (Bases 1 to 161)
Roebroek,A.J.M.
Direct Submission
Submitted (10-APR-2001) Center for Human Genetics, KU Leuven,
Herestraat 49, Leuven B-3000, Belgium
JOURNAL
FEATURES
source
1..161
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129P3/J"
/db_xref="taxon:10090"
11..151
/feature="Lrp1"
/number=54
exon

ORIGIN
Alignment Scores:
Pred. No.: 1.64e+03 Length: 161
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 10 Gaps: 0

US-10-030-194A-6 (1-6) x F369389S54 (1-161)

QY 1 GlyTyr***ValGlu 5
Db 17 GGGTACTCTGTGGAG 3

RESULT 33
BX663632/c 165 bp DNA linear STS 11-OCT-2003
LOCUS
DEFINITION Arabidopsis thaliana transposon insertion STS AT_5.12016, sequence
tagged site.
ACCESSION BX663632
VERSION BX663632.1 GI:37651991
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (Chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```



Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Walsh,S., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.

Unpublished

2 (bases 1 to 165)

Clarke,J.H.

Direct Submission

Submitted (09-OCT-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK

AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. 3 denotes a sequence derived from the 3' end of the transposon. 5 denotes a sequence derived from the 5' end of the transposon. BESRC GARNET, ATIS project

On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock code: NI02388.

Location/Qualifiers

1.165

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/variety="Landsberg erecta NASC stock code NW20"

/cultivar="Tn113Xtn25"

/db\_xref="taxon:3702"

/clone="AL021889"

1.165

/standard\_name="AT\_5.12016"

STS

ORIGIN

Alignment Scores:

Pred. No.: 1.69e+03 Length: 165

Score: 22.00 Matches: 4

Percent Similarity: 80.00% Conservative: 0

Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0

DB: 11 Gaps: 0

US-10-030-194A-6 (1-6) x BX663632 (1-165)

QY 1 GlyTyr\*\*\*ValGlu 5

|||||

DB 21 GCGTATAGTAGAGAA 7

RESULT 34

AR358585

LOCUS AR358585 166 bp DNA linear PAT 17-AUG-2003

DEFINITION Sequence 4703 from patent US 6593114.

ACCESSION AR358585

VERSION AR358585.1 GI:33764669

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 166)

AUTHORS Kunesch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and Rosen,C.A.

TITLE Staphylococcus aureus polynucleotides and sequences

JOURNAL Patent: US 6593114-A 4703 15-JUL-2003;

FEATURES

Location/Qualifiers

1.166

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 1.7e+03 Length: 166

Score: 22.00 Matches: 4

Percent Similarity: 80.00% Conservative: 0

Best Local Similarity: 80.00% Mismatches: 0

Query Match: 91.67% Indels: 0

DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x AR358585 (1-166)

QY 1 GlyTyr\*\*\*ValGlu 5

|||||

DB 13 GGTATTCTGTGAA 27

RESULT 35

CQ434907

LOCUS CQ434907 169 bp DNA linear PAT 30-JAN-2004

DEFINITION Sequence 667 from Patent WO0192523.

ACCESSION CQ434907

VERSION CQ434907.1 GI:41411669

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Shimkets,R.A. and Leach,M.D.

Human polynucleotides and polypeptides encoded thereby

Patent: WO 0192523-A 667 06-DEC-2001;

Curagen Corporation (US)

Location/Qualifiers

1.169

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1.73e+03 Length: 169

Score: 22.00 Matches: 4

Percent Similarity: 80.00% Conservative: 0

Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0

DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ434907 (1-169)

QY 1 GlyTyr\*\*\*ValGlu 5

|||||

DB 126 GGTATTTCAGTGGAG 140

RESULT 36

HS301D7/c

LOCUS HS301D7/c 172 bp DNA linear PRI 12-SEP-1996

DEFINITION H.sapiens chromosome 22 CpG island DNA genomic MseI fragment, clone 301d7, complete read.

ACCESSION Z79819

VERSION Z79819.1 GI:1531807

KEYWORDS Chromosome 22; CpG island.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.

Purification of CpG islands using a methylated DNA binding column

Nat. Genet. 6 (3), 236-244 (1994)

94282070

PUBMED 8012384

REFERENCE 2 (bases 1 to 172)

AUTHORS Clark,V.H., Cross,S.H., Simmen,M.W., Langford,C., Carter,N., Bickmore,W. and Bird,A.P.

Direct Submission

TITLE Submitted (02-SEP-1996) ICMB, University of Edinburgh, King's Buildings, Mayfield Rd, Edinburgh EH9 3JR, UK. E-mail contact: VCLARK@srv0.bio.ed.ac.uk

JOURNAL Vector: pGEM-5zf(-).

FEATURES

Location/Qualifiers

1.172

source

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="sorted chromosome 22"
/clone="301d7"
/cell_line="lymphoblastoid"
/clone_lib="CGI22.1"

```

ORIGIN

Alignment Scores:  
 Pred. No.: 1.76e+03 Length: 172  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservaive: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x HS301D7 (1-172)

QY 1 GlyTyr\*\*\*ValGlu 5  
 ||||| |||||  
 Db 19 GGGTATACGGTAGAG 5

RESULT 37

BX465325/c  
 LOCUS Arabidopsis thaliana transposon insertion STS SM\_3.36930, sequence tagged site.  
 DEFINITION BX465325 172 bp DNA linear STS 06-JUN-2003

ACCESSION BX465325.1 GI:30139476  
 VERSION STS; STS, sequence tagged site.  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B.,  
 Murphy,G., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 172)  
 AUTHORS Clarke,J.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-APR-2003) Clarke J.H., John Innes Centre, Colney  
 Lane, Norwich, NR4 7UJ, UK

COMMENT Ar denotes an activation tag dissociation transposon within a  
 single line, Et an enhancer trap dissociation transposon, GT a gene  
 trap dissociation transposon, MT a mis-expression enhancer trap  
 dissociation transposon, SM a defective suppressor mutator  
 transposon. \_3 denotes a sequence derived from the 3'end of the  
 transposon. \_5 denotes a sequence derived from the 5'end of the  
 transposon B5RC GARNET, ATIS project  
 On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock  
 code: N123641.

FEATURES

source Location/Qualifiers  
 1..172  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /variety="Columbia-0 NASC stock code N1092"  
 /db\_xref="taxon:3702"  
 /clone="AL162873"  
 /notes="Derived from superpool 23.16 NASC code N41063"  
 1..172  
 /standard\_name="SM\_3.36930"

STS

ORIGIN

Alignment Scores:  
 Pred. No.: 1.76e+03 Length: 172  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservaive: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 11 Gaps: 0

US-10-030-194A-6 (1-6) x BX465325 (1-172)

QY 1 GlyTyr\*\*\*ValGlu 5  
 ||||| |||||  
 Db 172 GGTATTTCGGTTGAA 158

RESULT 38

BX546436/c  
 LOCUS Arabidopsis thaliana transposon insertion STS SM\_3.41179, sequence tagged site.  
 DEFINITION BX546436 172 bp DNA linear STS 30-JUN-2003

ACCESSION BX546436.1 GI:32399317  
 VERSION STS; STS, sequence tagged site.  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
 AUTHORS Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Walsh,S.,  
 Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.

JOURNAL

REFERENCE 2 (bases 1 to 172)  
 AUTHORS Clarke,J.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-JUN-2003) Clarke J.H., John Innes Centre, Colney  
 Lane, Norwich, NR4 7UJ, UK

COMMENT Ar denotes an activation tag dissociation transposon within a  
 single line, Et an enhancer trap dissociation transposon, GT a gene  
 trap dissociation transposon, MT a mis-expression enhancer trap  
 dissociation transposon, SM a defective suppressor mutator  
 transposon. \_3 denotes a sequence derived from the 3'end of the  
 transposon. \_5 denotes a sequence derived from the 5'end of the  
 transposon B5RC GARNET, ATIS project  
 On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock  
 code: N127890.

FEATURES

source Location/Qualifiers  
 1..172  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /variety="Columbia-0 NASC stock code N1092"  
 /db\_xref="taxon:3702"  
 /clone="AL162873"  
 /notes="Derived from superpool 20.39 NASC code N40934"  
 1..172  
 /standard\_name="SM\_3.41179"

STS

ORIGIN

Alignment Scores:  
 Pred. No.: 1.76e+03 Length: 172  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservaive: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 11 Gaps: 0

US-10-030-194A-6 (1-6) x BX546436 (1-172)

QY 1 GlyTyr\*\*\*ValGlu 5  
 ||||| |||||  
 Db 172 GGTATTTCGGTTGAA 158

RESULT 39

BX465370/c  
 LOCUS Arabidopsis thaliana transposon insertion STS SM\_3.36959, sequence tagged site.  
 DEFINITION BX465370 175 bp DNA linear STS 06-JUN-2003

ACCESSION BX465370.1 GI:30139521  
 VERSION STS; STS, sequence tagged site.  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana

```

ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
AUTHORS    Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B.,
Murphy, G., Lingham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.
Unpublished
2 (bases 1 to 175)
JOURNAL    Submitted (23-APR-2003) Clarke J.H., John Innes Centre, Colney
AUTHORS    Lane, Norwich, NR4 7UJ, UK
TITLE      Direct Submission
JOURNAL    Submitted (23-APR-2003) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
COMMENT    At denotes an activation tag dissociation transposon within a
single line, Et an enhancer trap dissociation transposon, Gr a gene
trap dissociation transposon, Mt a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. 3 denotes a sequence derived from the 3'end of the
transposon. 5 denotes a sequence derived from the 5'end of the
transposon. BSRG GARNET, ATIS project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
code: N123670.
FEATURES   Location/Qualifiers
source     1..175
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /variety="Columbia-0 NASC stock code N1092"
            /db_xref="taxon:3702"
            /clone="AL162873"
            /note="Derived from superpool 23.18 NASC code N41065"
            /standard_name="SM_3.36959"
ST          1..175
            /standard_name="SM_3.36959"
ORIGIN
Alignment Scores: 1.79e+03 Length: 175
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67% Gaps: 0
DB: 11
US-10-030-194A-6 (1-6) x BX465370 (1-175)
Qy 1 GlyTyr***ValGlu 5
Db 172 GGTATTTCGGTTGAA 158
RESULT 40
BX465345/c
LOCUS     BX465345
DEFINITION Arabidopsis thaliana transposon insertion STS SM_3.36964, sequence
tagged site.
ACCESSION BX465345
VERSION   BX465345.1 GI:30139496
KEYWORDS  STS; STS, sequence tagged site.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
AUTHORS    Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B.,
Murphy, G., Lingham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.
Unpublished
2 (bases 1 to 176)
JOURNAL    Submitted (05-JUL-2002) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
COMMENT    At denotes an activation tag dissociation transposon within a
single line, Et an enhancer trap dissociation transposon, Gr a gene
trap dissociation transposon, Mt a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. 3 denotes a sequence derived from the 3'end of the
transposon. 5 denotes a sequence derived from the 5'end of the
transposon. BSRG GARNET, ATIS project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
code: N108194.
FEATURES   Location/Qualifiers
source     1..184
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /variety="Columbia-0 NASC stock code N1092"
            /db_xref="taxon:3702"
            /clone="AL162873"
            /note="Derived from superpool 20.48 NASC code N40943"
            /standard_name="SM_3.19046"
ST          1..184
            /standard_name="SM_3.19046"

```

```

FEATURES   Location/Qualifiers
source     1..176
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /variety="Columbia-0 NASC stock code N1092"
            /db_xref="taxon:3702"
            /clone="AL162873"
            /note="Derived from superpool 23.18 NASC code N41065"
            /standard_name="SM_3.36964"
ST          1..176
            /standard_name="SM_3.36964"
ORIGIN
Alignment Scores: 1.8e+03 Length: 176
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67% Gaps: 0
DB: 11
US-10-030-194A-6 (1-6) x BX465345 (1-176)
Qy 1 GlyTyr***ValGlu 5
Db 172 GGTATTTCGGTTGAA 158
RESULT 41
AL831542/c
LOCUS     AL831542
DEFINITION Arabidopsis thaliana transposon insertion STS SM_3.19046, sequence
tagged site.
ACCESSION AL831542
VERSION   AL831542.1 GI:21713189
KEYWORDS  STS; STS, sequence tagged site.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
AUTHORS    Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B.,
Murphy, G., Lingham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.
Unpublished
2 (bases 1 to 184)
JOURNAL    Submitted (05-JUL-2002) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
COMMENT    At denotes an activation tag dissociation transposon within a
single line, Et an enhancer trap dissociation transposon, Gr a gene
trap dissociation transposon, Mt a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. 3 denotes a sequence derived from the 3'end of the
transposon. 5 denotes a sequence derived from the 5'end of the
transposon. BSRG GARNET, ATIS project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
code: N108194.
FEATURES   Location/Qualifiers
source     1..184
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /variety="Columbia-0 NASC stock code N1092"
            /db_xref="taxon:3702"
            /clone="AL162873"
            /note="Derived from superpool 20.48 NASC code N40943"
            /standard_name="SM_3.19046"
ST          1..184
            /standard_name="SM_3.19046"

```

RESULT 43		BT011699		186 bp mRNA linear		PLN 04-MAR-2004	
LOCUS		BT011699		Arabidopsis thaliana At5g55896 gene, complete cds.			
DEFINITION		BT011699					
ACCESSION		BT011699.1		GI:44917474			
VERSION		BT011699.1		FLI CDNA.			
KEYWORDS				Arabidopsis thaliana (thale cress)			
SOURCE				Arabidopsis thaliana			
ORGANISM				Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi			
REFERENCE				1 (bases 1 to 186)			
AUTHORS				Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Carninci,P., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Narusaka,M., Sakurai,T., Satou,M., Seki,M., Shinozaki,K. and Ecker,J.R.			
TITLE				Arabidopsis ORF clones			
JOURNAL				Unpublished			
REFERENCE				2 (bases 1 to 186)			
AUTHORS				Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Carninci,P., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Narusaka,M., Sakurai,T., Satou,M., Seki,M., Shinozaki,K. and Ecker,J.R.			
TITLE				Direct Submission			
JOURNAL				Submitted (04-MAR-2004) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA			
COMMENT				RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.			
FEATURES		source		1..186		Location/Qualifiers	
				/organism="Arabidopsis thaliana"			
				/mol_type="mRNA"			
				/db_xref="taxon:3702"			
				/chromosome="5"			
				/clone="U65938"			
				/ecotype="Columbia"			
CDS				1..186		/note="unknown protein"	
				/codon_start=1			
				/product="At5g55896"			
				/protein_id="AA549062.1"			
				/db_xref="GI:44917475"			
				/translation="MMELKRLSSLFGRIGRLVGRMATQLSFLEILGALLVRKLLPSMS SLILGSSLSGMPPLWF"			
ORIGIN							
Alignment Scores:							
Pred. No.:		1.91e+03		Length:		186	
Score:		22.00		Matches:		4	
Percent Similarity:		80.00%		Conservative:		0	
Best Local Similarity:		80.00%		Mismatch:		1	
Query Match:		91.67%		Indels:		0	
DB:		11		Gaps:		0	
US-10-030-194A-6 (1-6) x AL831542 (1-184)							
QY		1 GlyTyr***ValGlu 5					
DB		172 GGTATTTCGGTTGAA 158					
RESULT 42							
BX664061/c							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
COMMENT							
FEATURES		source		1..184		Location/Qualifiers	
				/organism="Arabidopsis thaliana"			
				/mol_type="genomic DNA"			
				/variety="Columbia-0 NASC stock code N1092"			
				/db_xref="taxon:3702"			
				/clone="AL162873"			
				/notes="Derived from superpool 23.05 NASC code N41052"			
STS				1..184		/standard_name="SM_3.16812"	
ORIGIN							
Alignment Scores:							
Pred. No.:		1.89e+03		Length:		184	
Score:		22.00		Matches:		4	
Percent Similarity:		80.00%		Conservative:		0	
Best Local Similarity:		80.00%		Mismatch:		1	
Query Match:		91.67%		Indels:		0	
DB:		11		Gaps:		0	
US-10-030-194A-6 (1-6) x BX664061 (1-184)							
QY		1 GlyTyr***ValGlu 5					
DB		172 GGTATTTCGGTTGAA 158					

```

Db          |||||  |||||  |||||
60 GGCTACTCAGTTGAG 74

HUMINSRA02      188 bp      DNA      linear      PRI 05-FEB-1996
LOCUS           Human insulin receptor gene allele A, exon 18.
DEFINITION      M27197
ACCESSION       M27197.1 GI:186466
VERSION          insulin receptor.
KEYWORDS         2 of 2
SEGMENT          Homo sapiens (human)
SOURCE           Homo sapiens
ORGANISM         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS          1 (bases 1 to 188)
                  Taïra, M., Taïra, M., Hashimoto, N., Shimada, F., Suzuki, Y.,
                  Kanatsuka, A., Nakamura, F., Ebina, Y., Tatibana, M., Makino, H. and
                  Yoshida, S.
TITLE            Human diabetes associated with a deletion of the tyrosine kinase
                  domain of the insulin receptor
JOURNAL          Science 245 (4913), 63-66 (1989)
MEDLINE          89298408
PubMed          2344997
COMMENT          Original source text: Human (patient 1, haplotype A2) white
                  bloodcell DNA.
                  Draft entry and computer-readable sequence for [1] kindly provided
                  by M.Taira, 02-SEP-1989.
FEATURES         Location/Qualifiers
source           1..188
                 /organism="Homo sapiens"
                 /mol_type="genomic DNA"
                 /db_xref="taxon:9606"
                 /map="19p13.3-p13.2"
gene            join(M27195.1:40..349,1..122)
                 /gene="INSR"
                 join(M27195.1:<40..284,12..>122)
CDS             /gene="INSR"
                 /notes="insulin receptor"
                 /codon_start=3
                 /protein_id="AA886791.1"
                 /db_xref="GI:186468"
                 /translations="FPCSYYVPDWEVSREKITLLRELGGQSGFMVYEGNARDIIGKE
AETRVAVKTVNESASLRERIEFLNEASVMKGFTCHHVRLLGVSQKQPTLVVWELMA"
                 HGDLKSYLRSLRPEAE"
intron          <1..11
                 /gene="INSR"
                 /notes="INSR intron Q"
exon            12..122
                 /gene="INSR"
                 /notes="insulin receptor; G00-119-352"
                 /number=18
intron          123..>188
                 /notes="INSR intron R"
ORIGIN          About 5 kb after segment 1.

Alignment Scores:
Pred. No.:      1.93e+03      Length:      188
Score:          22.00         Matches:      4
Percent Similarity: 80.00%    Conservative: 0
Best Local Similarity: 80.00% Mismatches:      1
Query Match:    91.67%       Indels:       0
DB:             Gaps:        0

US-10-030-194A-6 (1-6) x HUMINSRA02 (1-188)

Qy            1 GlyTyr***ValGlu 5
              |||||
Db            148 GGCTACTCGGTGGAG 162

RESULT 45
LOCUS          CQ059104/c
FEATURES       CQ059104

```

```
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ183536 (1-189)

QY 1 GlyTyr***ValGlu 5
Db 138 GGCTATGCTGTAGAA 124

RESULT 47
CQ269587/c 189 bp DNA linear PAT 23-JAN-2004
LOCUS
DEFINITION Sequence 17848 from Patent WO0157277.
ACCESSION CQ269587
VERSION CQ269587.1 GI:41242191
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human fetal liver
JOURNAL Patent: WO 0357277-A 17848 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source Location/Qualifiers
1..189
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009576.2-EXPRESSED IN FETAL LIVER, SIGNAL
= 1.1-EST HUMAN HIT: BE617547.1, EVALUE 2.60e+00-SWISSPROT
HIT: Q58850, EVALUE 2.90e+00-NT HIT: U63374.1, EVALUE
9.50e-02"
182

ORIGIN
unsure
Alignment Scores: 1.94e+03 Length: 189
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67% Gaps: 0
DB: 6

US-10-030-194A-6 (1-6) x CQ269587 (1-189)

QY 1 GlyTyr***ValGlu 5
Db 138 GGCTATGCTGTAGAA 124

RESULT 48
CQ306668/c 189 bp DNA linear PAT 23-JAN-2004
LOCUS
DEFINITION Sequence 17773 from Patent WO0186003.
ACCESSION CQ306668
VERSION CQ306668.1 GI:41267245
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human lung
JOURNAL Patent: WO 0186003-A 17773 15-NOV-2001;
Aeomica, Inc. (US)
FEATURES
source Location/Qualifiers
1..189
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009576.2-EXPRESSED IN FETAL LIVER, SIGNAL
= 1.1-EST HUMAN HIT: BE617547.1, EVALUE 2.60e+00-SWISSPROT
HIT: Q58850, EVALUE 2.90e+00-NT HIT: U63374.1, EVALUE
9.50e-02"
182

ORIGIN
unsure
Alignment Scores: 1.94e+03 Length: 189
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67% Gaps: 0
DB: 6

US-10-030-194A-6 (1-6) x CQ269587 (1-189)

QY 1 GlyTyr***ValGlu 5
Db 138 GGCTATGCTGTAGAA 124

RESULT 49
AR206039/c 193 bp DNA linear PAT 20-JUN-2002
LOCUS
DEFINITION Sequence 53 from patent US 6372211.
ACCESSION AR206039
VERSION AR206039.1 GI:21504522
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 193)
AUTHORS Isaac,B.G., Greenplate,J.T., Purcell,J.P. and Romano,C.P.
TITLE Methods and compositions for controlling insects
JOURNAL Patent: US 6372211-A 53 16-APR-2002;
FEATURES
source Location/Qualifiers
1..193
/organism="unknown"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009576.2-EXPRESSED IN LUNG, SIGNAL =
10-EST HUMAN HIT: BE617547.1, EVALUE 2.60e+00-SWISSPROT
HIT: Q58850, EVALUE 2.90e+00-NT HIT: U63374.1, EVALUE
9.50e-02"
182

ORIGIN
unsure
Alignment Scores: 1.99e+03 Length: 193
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67% Gaps: 0
DB: 6

US-10-030-194A-6 (1-6) x AR206039 (1-193)

QY 1 GlyTyr***ValGlu 5
Db 147 GGCTACACGGTTGAG 133

RESULT 50
HS39CSF 194 bp DNA linear PRI 17-OCT-1995
LOCUS
DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 39c5, forward
read cpg39c5.ftie.
ACCESSION Z55399
VERSION Z55399.1 GI:1021440
KEYWORDS CpG island; genomic MseI fragment.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009576.2-EXPRESSED IN LUNG, SIGNAL =
10-EST HUMAN HIT: BE617547.1, EVALUE 2.60e+00-SWISSPROT
HIT: Q58850, EVALUE 2.90e+00-NT HIT: U63374.1, EVALUE
9.50e-02"
182

ORIGIN
unsure
Alignment Scores: 1.94e+03 Length: 189
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67% Gaps: 0
DB: 6

US-10-030-194A-6 (1-6) x CQ306668 (1-189)

QY 1 GlyTyr***ValGlu 5
Db 138 GGCTATGCTGTAGAA 124

RESULT 49
AR206039/c 193 bp DNA linear PAT 20-JUN-2002
LOCUS
DEFINITION Sequence 53 from patent US 6372211.
ACCESSION AR206039
VERSION AR206039.1 GI:21504522
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 193)
AUTHORS Isaac,B.G., Greenplate,J.T., Purcell,J.P. and Romano,C.P.
TITLE Methods and compositions for controlling insects
JOURNAL Patent: US 6372211-A 53 16-APR-2002;
FEATURES
source Location/Qualifiers
1..193
/organism="unknown"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009576.2-EXPRESSED IN LUNG, SIGNAL =
10-EST HUMAN HIT: BE617547.1, EVALUE 2.60e+00-SWISSPROT
HIT: Q58850, EVALUE 2.90e+00-NT HIT: U63374.1, EVALUE
9.50e-02"
182

ORIGIN
unsure
Alignment Scores: 1.99e+03 Length: 193
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Percent Similarity: 80.00% Mismatches: 1
Best Local Similarity: 80.00% Indels: 0
Query Match: 91.67% Gaps: 0
DB: 6

US-10-030-194A-6 (1-6) x AR206039 (1-193)

QY 1 GlyTyr***ValGlu 5
Db 147 GGCTACACGGTTGAG 133

RESULT 50
HS39CSF 194 bp DNA linear PRI 17-OCT-1995
LOCUS
DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 39c5, forward
read cpg39c5.ftie.
ACCESSION Z55399
VERSION Z55399.1 GI:1021440
KEYWORDS CpG island; genomic MseI fragment.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
```

```

MEDLINE 94282070
PUBMED 8012384
REFERENCE 2 (bases 1 to 194)
AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: bihelp@hgmp.mrc.ac.uk.

FEATURES
source
1..194
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="39c5"
/sex="male"
/tissue_type="blood"
/clone_lib="CGI-1"
/dev_stage="adult"

ORIGIN
Alignment Scores:
Pred. No.: 2e+03 Length: 194
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x HS39C5F (1-194)

Qy 1 GlyTyr***ValGlu 5
Db 94 GGTACTACTGTGTGAA 108

RESULT 51
LOCUS CQ713883 200 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 58809 from Patent WO02070737.
ACCESSION CQ713883
VERSION CQ713883.1 GI:42274740
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 58809 12-SEP-2002;
Chondrogene Inc. (CA)

FEATURES
source
1..200
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.06e+03 Length: 200
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CQ713883 (1-200)

Qy 1 GlyTyr***ValGlu 5
|||||

```

```

Db 63 GGATATTCTGTAGAA 77

RESULT 52
LOCUS BVI75451 201 bp DNA linear STS 10-JUN-2004
DEFINITION sqnm79864 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
tagged site.
ACCESSION BVI75451
VERSION BVI75451.1 GI:48011037
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 201)
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
Regions
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018
Fax: 18582029020
Email: abraun@sequenom.com
Primer A: No primer sequence submitted
Primer B: No primer sequence submitted
STS size: 201.
Location/Qualifiers
1..201
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human DNA (Sequenom)"
<1..>201

ORIGIN
Alignment Scores:
Pred. No.: 2.07e+03 Length: 201
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-6 (1-6) x BVI75451 (1-201)

Qy 1 GlyTyr***ValGlu 5
|||||
Db 31 GGGTATGCTGTGGAG 45

RESULT 53
LOCUS BV204514 201 bp DNA linear STS 10-JUN-2004
DEFINITION sqnm215472 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
tagged site.
ACCESSION BV204514
VERSION BV204514.1 GI:48174192
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 201)
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
Regions
JOURNAL Genome Res. (2004) In press
COMMENT

```

Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 201.

FEATURES  
source  
1. .201  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
<1. .>201

ORIGIN  
Alignment Scores: Length: 201  
Pred. No.: 2.07e+03 Matches: 4  
Score: 22.00 Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67%  
DB: 11 Gaps: 0

US-10-030-194A-6 (1-6) x BV204514 (1-201)

QY 1 GlyTyr\*\*\*ValGlu 5  
||||| |||||

Db 45 GGATACACAGTGGAG 59

RESULT 54  
AR269722/c  
LOCUS AR269722 202 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 285 from patent US 6500938.  
ACCESSION AR269722  
VERSION AR269722.1 GI:29700954  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 202)  
AUTHORS Au-Young, J. and Seilhamer, J. J.  
TITLE Composition for the detection of signaling pathway gene expression  
JOURNAL Patent: US 6500938-A 285 31-DEC-2002;  
FEATURES  
source  
1. .202  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores: Length: 202  
Pred. No.: 2.08e+03 Matches: 4  
Score: 22.00 Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67%  
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x AR269722 (1-202)

QY 1 GlyTyr\*\*\*ValGlu 5  
||||| |||||

Db 86 GGCTATACAGTGGAG 72

RESULT 55  
BX284270 204 bp DNA linear STS 11-JUN-2003  
LOCUS BX284270  
DEFINITION Arabidopsis thaliana transposon insertion STS SM\_3.23777, sequence tagged site.  
ACCESSION BX284270

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BX284270.1 GI:28803653  
STS, sequence tagged site.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE  
AUTHORS

1  
Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B.,  
Murphy, G., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 204)

AUTHORS

Clarke, J.H.

TITLE

Direct Submission

JOURNAL

Submitted (03-MAR-2003) Clarke J.H., John Innes Centre, Colney  
Lane, Norwich, NR4 7UJ, UK

COMMENT

AT denotes an activation tag dissociation transposon within a  
single line, ET an enhancer trap dissociation transposon, GT a gene  
trap dissociation transposon, MT a mis-expression enhancer trap  
dissociation transposon, SM a defective suppressor mutator  
transposon. \_3 denotes a sequence derived from the 3' end of the  
transposon. \_5 denotes a sequence derived from the 5' end of the  
transposon. BBSRC GARNET, ATIS project  
On-line seed stock requests: <http://nasc.nott.ac.uk/NASC> stock  
code: N112233.

FEATURES  
source

1. .204  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia-0 NASC stock code N1092"  
/db\_xref="taxon:3702"  
/clone="AC012563"  
/note="Derived from superpool 7.18 NASC code N40313"  
1. .204  
/standard\_name="SM\_3.23777"

STS

ORIGIN

Alignment Scores: Length: 204  
Pred. No.: 2.1e+03 Matches: 4  
Score: 22.00 Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67%  
DB: 11 Gaps: 0

US-10-030-194A-6 (1-6) x BX284270 (1-204)

QY 1 GlyTyr\*\*\*ValGlu 5  
||||| |||||

Db 125 GGTATTTCAGTCGAA 139

RESULT 56

AX919636

LOCUS

AX919636 206 bp DNA linear PAT 18-DEC-2003

DEFINITION

Sequence 35499 from Patent EP1033401.

ACCESSION

AX919636

VERSION

AX919636.1 GI:40213425

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.

TITLE

Expressed sequence tags and encoded human proteins

JOURNAL

Patent: EP 1033401-A 35499 06-SEP-2000;

FEATURES

source

1. .206

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN



```

Alignment Scores:
Pred. No.: 2.12e+03 Length: 206
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x AX919636 (1-206)

Oy 1 GlyTyr***ValGlu 5
||| ||| ||| |||
Db 154 GGGTATGCTGTTGAG 168

RESULT 57
BD055169 206 bp DNA linear PAT 27-AUG-2002
LOCUS Sequence tag and encoded human protein.
ACCESSION BD055169
VERSION BD055169.1 GI:22600775
KEYWORDS JP 2001269182-A/31415.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 206)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 31415 02-OCT-2001;
GENESEN

COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/31415
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PJ JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/03, C12P21/08, C12Q1/68 //G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers.
FEATURES
source Location/Qualifiers
1..206
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.12e+03 Length: 206
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x BD055169 (1-206)

Oy 1 GlyTyr***ValGlu 5
||| ||| ||| |||
Db 154 GGGTATGCTGTTGAG 168

RESULT 58
SYNDHFRII 220 bp DNA linear SYN 27-APR-1993
LOCUS Synthetic mini type II dihydrofolate reductase, complete cds.
DEFINITION M20407
ACCESSION M20407.1 GI:208215
KEYWORDS dihydrofolate reductase.
SOURCE synthetic construct

ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 220)
AUTHORS Vermersch,P.S. and Bennett,G.N.
TITLE Synthesis and expression of a gene for a mini type II dihydrofolate
reductase
JOURNAL DNA 8, 243-251 (1988)
COMMENT Original source text: Synthetic DNA.
FEATURES
source Location/Qualifiers
1..220
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
8..214
/notes="dihydrofolate reductase"
/codon_start=1
/transl_table=11
/protein_id="AAA72656.1"
/db_xref="GI:208216"
/translation="MASALPSLPSSATFGLGDRVRKSGAANQGVGVGYCTKLTPEG
YAVESHPGVSQIYPVALERVA"
ORIGIN 1 bp upstream of BamHI site.

Alignment Scores:
Pred. No.: 2.28e+03 Length: 220
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 12 Gaps: 0

US-10-030-194A-6 (1-6) x SYNDHFRII (1-220)

Oy 1 GlyTyr***ValGlu 5
||| ||| ||| |||
Db 137 GGCTACGGGTTGAG 151

RESULT 59
F260679S01 221 bp DNA linear PRI 02-JAN-2003
LOCUS Homo sapiens acyl-CoA dehydrogenase 8 (ACAD8) gene, exon 1.
DEFINITION AP260679
ACCESSION AP260679.1 GI:9739139
VERSION AP260679.1
KEYWORDS 1 of 11
SEGMENT Homo sapiens (human)
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 221)
AUTHORS Nguyen,T.V., Andresen,B.S., Corydon,T.J., Ghisla,S., Abd-El
Razik,N.J. and Vockley,J.
TITLE Identification of isobutyryl-CoA dehydrogenase and its deficiency
in humans
JOURNAL Mol. Genet. Metab. 77 (1-2), 68-79 (2002)
MEDLINE 22247156
PUBMED 12359132
REFERENCE 2 (bases 1 to 221)
AUTHORS Andresen,B.S.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2000) Research Unit for Molecular Medicine and
Institute of Human Genetics, Aarhus University, Brendstrupgaardsvej
100, Aarhus N. 8200 N., Denmark
FEATURES
source Location/Qualifiers
1..221
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q25"
<1..166
/gene="ACAD8"
exon

```

```
ORIGIN
Alignment Scores:
Pred. No.: 2.36e+03 Length: 228
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservatives: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x AR395391 (1-228)
QY 1 GlyTyr***ValGlu 5
Db 131 GGCTATTGAGTGGAA 145

RESULT 62
BX323626/c
LOCUS 237 bp DNA linear STS 10-JUN-2003
DEFINITION Arabidopsis thaliana transposon insertion STS SM_3.37256, sequence tagged site.
ACCESSION BX323626
VERSION BX323626.1 GI:29787833
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1
AUTHORS Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Murphy,G., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 237)
AUTHORS Clarke,J.H.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK
COMMENT AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. _3 denotes a sequence derived from the 3' end of the transposon, _5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ATIS project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock code: N123967.
FEATURES
source Location/Qualifiers
1..237
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AL162873"
/note="Derived from superpool 23.06 NASC code N41053"
1..237
/standard_name="SM_3.37256"

STS
ORIGIN
Alignment Scores:
Pred. No.: 2.46e+03 Length: 237
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservatives: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-6 (1-6) x BX323626 (1-237)
QY 1 GlyTyr***ValGlu 5
Db 172 GGTTATTCGTTGAA 158

RESULT 63
G02926/c
```

```
ORIGIN
Alignment Scores:
Pred. No.: 2.29e+03 Length: 221
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservatives: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x F260679S01 (1-221)
QY 1 GlyTyr***ValGlu 5
Db 54 GGCTATGCTGTGGAG 68

RESULT 60
AR396502/c
LOCUS 225 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2517 from patent US 6617156.
ACCESSION AR396502
VERSION AR396502.1 GI:40126061
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 225)
AUTHORS Doucette-Stamm,L.A. and Bush,D.
TITLE Nucleic acid and amino acid sequences relating to Enterococcus faecalis for diagnostics and therapeutics
JOURNAL Patent: US 6617156-A 2517 09-SEP-2003;
FEATURES
source Location/Qualifiers
1..225
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 2.33e+03 Length: 225
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservatives: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x AR396502 (1-225)
QY 1 GlyTyr***ValGlu 5
Db 118 GGTTATAGTGTGAA 104

RESULT 61
AR395391
LOCUS 228 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1406 from patent US 6617156.
ACCESSION AR395391
VERSION AR395391.1 GI:40123942
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 228)
AUTHORS Doucette-Stamm,L.A. and Bush,D.
TITLE Nucleic acid and amino acid sequences relating to Enterococcus faecalis for diagnostics and therapeutics
JOURNAL Patent: US 6617156-A 1406 09-SEP-2003;
FEATURES
source Location/Qualifiers
1..228
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
```

LOCUS G02926 237 bp DNA linear STS 19-OCT-1995  
 DEFINITION human STS WI-2317, sequence tagged site.  
 ACCESSION G02926  
 VERSION G02926.1 GI:719884  
 KEYWORDS STS; STS sequence; primer; sequence tagged site.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 237)  
 AUTHORS Hudson,T.  
 TITLE Whitehead Institute/MIT Center for Genome Research; Random Genome  
 JOURNAL Wide STS  
 REFERENCE 2 (bases 1 to 237)  
 AUTHORS Hudson,T.  
 TITLE Whitehead Institute/MIT Center for Genome Research; Physically  
 Mapped STSs  
 JOURNAL Unpublished (1995)  
 COMMENT GDB DSEG: D6S1074  
 Contact: Thomas Hudson  
 Whitehead Institute/MIT Center for Genome Research  
 Whitehead Institute for Biomedical Research  
 9 Cambridge Center, Cambridge MA 02142 USA  
 Tel: 617 252 1900  
 Fax: 617 252 1902  
 Email: thudson@genome.wi.mit.edu  
 Primer A: TCCCTTCTCTACTGTAATACCCAC  
 Primer B: CATGTGTCATGAGACTGGG  
 STS size: 104  
 PCR Profile:  
 Presoak:  
 Denaturation:  
 Annealing: 56 degrees C  
 Polymerization:  
 PCR Cycles: 35  
 Thermal Cycler:  
 Protocol:  
 Template: 10 ng  
 Primer: each 5 pM  
 dNTPs: each 4 mM  
 Taq Polymerase: 0.025 units/ul  
 Total Vol: 20 ul  
 Buffer:  
 MgCl2: 1.5 mM  
 KCl: 50 mM  
 Tris-HCL: 10 mM  
 pH: 9.3.  
 FEATURES  
 source  
 1..237  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /map="764\_H\_9; 790\_G\_7; 899\_G\_1; 711\_F\_12; 767\_G\_12"  
 99..202  
 99..122  
 primer\_bind  
 primer\_bind complement(183..202)  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.46e+03 Length: 237  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 11 Gaps: 0  
 US-10-030-194A-6 (1-6) x G02926 (1-237)  
 Qy 1 GlyTyr\*\*\*ValGlu 5

||||| |||||  
 Db 119 GGTATTTCAGTAGAG 105  
 RESULT 64  
 G28679  
 LOCUS G28679 238 bp DNA linear STS 28-SEP-1998  
 DEFINITION sWSS3790 Eric D. Green Homo sapiens STS genomic, sequence tagged  
 site.  
 ACCESSION G28679  
 VERSION G28679.1 GI:1408115  
 KEYWORDS STS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 238)  
 AUTHORS Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,  
 Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S.,  
 Leckie,M.P. and Green,E.D.  
 TITLE A collection of 1814 human chromosome 7-specific STSs  
 JOURNAL Genome Res. 7 (1), 59-64 (1997)  
 MEDLINE 97189344  
 PUBMED 9037602  
 REFERENCE 2 (bases 1 to 238)  
 AUTHORS Green,E.D.  
 TITLE Human chromosome 7 STSs (1997)  
 JOURNAL Unpublished (1997)  
 COMMENT GDB: GDB:3755015  
 Contact: Eric D. Green  
 Genome Technology Branch  
 National Human Genome Research Institute/NIH  
 49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892  
 Tel: 3014020201  
 Fax: 3014024735  
 Email: egreen@nhgri.nih.gov  
 Primer A: GTTCAATTTCTCTGTACTC  
 Primer B: CTTTGTCATTTCCAGTCC  
 STS size: 61  
 PCR Profile:  
 Presoak: 0 degrees C for 0.00 minute(s)  
 Denaturation: 92 degrees C for 0.17 minute(s)  
 Annealing: 55 degrees C for 1.00 minute(s)  
 Polymerization: 72 degrees C for 1.00 minute(s)  
 PCR Cycles: 35  
 Thermal Cycler: PerkinElmer 9600  
 Protocol:  
 Template: 30-100 ng  
 Primer: each 1 uM  
 dNTPs: each 200 uM  
 Taq Polymerase: 0.05 units/ul  
 Total Vol: 10 ul  
 Buffer:  
 MgCl2: 1.5 mM  
 KCl: 100 mM  
 Tris-HCL: 10 mM  
 NH4Cl: 5 mM  
 pH: 8.6  
 The sequence for this STS was derived from a single sequencing  
 read. For additional information about the NHGRI chromosome 7  
 mapping project, see <http://www.nhgri.nih.gov/DIR/STB/CHR7>. Also  
 see Genomics 11:548-64 (1991) [MUID=92128937].  
 FEATURES  
 source  
 1..238  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /map="7"  
 /clone\_lib="Eric D. Green"  
 2..62  
 primer\_bind  
 2..20

```

    primer_bind      complement(45..62)
ORIGIN
Alignment Scores:      Length:      238
Pred. No.:             2.47e+03
Score:                 22.00
Matches:               4
Percent Similarity:    80.00%
Conservative:         0
Best Local Similarity: 80.00%
Mismatch:             1
Query Match:          91.67%
DB:                   0
Gaps:                 0

US-10-030-194A-6 (1-6) x G28679 (1-238)

Qy      1 GlyTyr***ValGlu 5
Db      125 GGCTACACAGTAGAA 139

RESULT 65
AF483760/c      240 bp      DNA      linear      MAM 14-OCT-2002
LOCUS           Bos taurus clone UMN0406 microsatellite sequence.
DEFINITION      AF483760
ACCESSION       AF483760.1 GI:23954457
VERSION         AF483760.1
KEYWORDS        Bos taurus (cow)
SOURCE          Bos taurus
ORGANISM        Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE       1 (bases 1 to 240)
AUTHORS        Liu,W.-S., Mariani,P., Beattie,C.W., Alexander,L.J. and Ponce de
                Leon,F.A.
TITLE          A radiation hybrid map for the bovine Y chromosome
JOURNAL        Unpublished
REFERENCE       2 (bases 1 to 240)
AUTHORS        Liu,W.-S., Mariani,P., Beattie,C.W., Alexander,L.J. and Ponce de
                Leon,F.A.
TITLE          Direct Submission
JOURNAL        Submitted (13-FEB-2002) Animal Science, University of Minnesota,
                305 Haacker Hall, 1364 Eckles Ave, St. Paul, MN 55108, USA
FEATURES        Location/Qualifiers
                source
                1..240
                /organism="Bos taurus"
                /mol_type="genomic DNA"
                /db_xref="taxon:9913"
                /chromosome="Y"
                /map="Y-specific region"
                /clone="UMN0406"
                /note="microsatellite"
                /rpt_type=tandem

repeat_region      1..240
                    /note="microsatellite"

ORIGIN
Alignment Scores:      Length:      240
Pred. No.:             2.49e+03
Score:                 22.00
Matches:               4
Percent Similarity:    80.00%
Conservative:         0
Best Local Similarity: 80.00%
Mismatch:             1
Query Match:          91.67%
DB:                   0
Gaps:                 0

US-10-030-194A-6 (1-6) x AF483760 (1-240)

Qy      1 GlyTyr***ValGlu 5
Db      15 GGATATTCGTAGAA 1

RESULT 66
AF549480
LOCUS           Coccotrypes carpophagus clone NTA-5-D10 microsatellite sequence.
DEFINITION      AF549480
ACCESSION       AF549480.1 GI:24415352
VERSION         AF549480.1

```

```

KEYWORDS        Coccotrypes carpophagus
SOURCE          Coccotrypes carpophagus
ORGANISM        Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
                Phytophaga; Scolytidae; Coccotrypes.
REFERENCE       1 (bases 1 to 241)
AUTHORS        Berg,P.R., Dawson,D.A., Kirkendall,L.R. and Burke,T.
TITLE          Isolation and characterization of microsatellite loci in two
                inbreeding bark beetle species, Coccotrypes
JOURNAL        Unpublished
REFERENCE       2 (bases 1 to 241)
AUTHORS        Berg,P.R. and Dawson,D.A.
TITLE          Direct Submission
JOURNAL        Submitted (03-OCT-2002) Department of Animal and Plant Sciences,
                University of Sheffield, Alfred Denny Building, Western Bank,
                Sheffield S10 2TN, UK
FEATURES        Location/Qualifiers
                source
                1..241
                /organism="Coccotrypes carpophagus"
                /mol_type="genomic DNA"
                /db_xref="taxon:123997"
                /clone="NTA-5-D10; CC17"
                /note="microsatellite"
                /rpt_type=tandem

repeat_region      1..241
                    /note="microsatellite"

ORIGIN
Alignment Scores:      Length:      241
Pred. No.:             2.5e+03
Score:                 22.00
Matches:               4
Percent Similarity:    80.00%
Conservative:         0
Best Local Similarity: 80.00%
Mismatch:             1
Query Match:          91.67%
DB:                   0
Gaps:                 0

US-10-030-194A-6 (1-6) x AF549480 (1-241)

Qy      1 GlyTyr***ValGlu 5
Db      155 GGATATTCGGTCGAG 169

RESULT 67
BV088606/c      242 bp      DNA      linear      STS 15-OCT-2003
LOCUS           RPAMMSEQ0000490 Roche Palo Alto Mus musculus STS genomic, sequence
DEFINITION      BV088606
ACCESSION       BV088606.1 GI:37666085
VERSION         BV088606
KEYWORDS        Mus musculus (house mouse)
SOURCE          Mus musculus
ORGANISM        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1 (bases 1 to 242)
AUTHORS        Usuka,J., Liao,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
                McPherson,J.D., Foernzier,D. and Peltz,G.
TITLE          Mus musculus SNPs
JOURNAL        Unpublished (2003)
COMMENT         Contact: Jonathan Usuka
                Roche Palo Alto Genetics and Genomics Department
                3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
                Tel: 6508555807
                Email: Jonathan.Usuka@roche.com
                Primer A: No primer submitted
                Primer B: No primer submitted.
                Location/Qualifiers
                1..242
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"

```

```

/map="8-2714-2951-CAA01135921.1.1.1.4800"
/clone_lib="Roche Palo Alto"
/note="SNPs developed from assay sequences derived from 15
different strains-of mice (as of October 1, 2003). Those
strains include A/J, A/HeJ, ~129/Sv, AKR/J, B10 D2-H2/cSnJ,
BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, -CAST/Ei, DBA/2J,
MRL/MpJ, NZB/BinJ, NZW/LacJ, SPRET/Ei."
STS
ORIGIN
Alignment Scores:
Pred. No.: 2.51e+03 Length: 242
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 11 Gaps: 0

US-10-030-194A-6 (1-6) x BV088606 (1-242)

Qy 1 GlyTyr***ValGlu 5
||||| |||||
Db 221 GGCTACAGTGTGGAA 207

RESULT 68
AL823809/c
LOCUS
DEFINITION
AL823809 243 bp DNA linear STS 10-JUN-2003
Arabidopsis thaliana transposon insertion STS SM_3.16284, sequence
tagged site.
ACCESSION
AL823809.1 GI:21701281
VERSION
STS; STS, sequence tagged site.
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS
Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B.,
Murphy,G., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.
Unpublished
2 (bases 1 to 243)
JOURNAL
Submitted (04-JUL-2002) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
COMMENT
AT denotes an activation tag dissociation transposon within a
single line, Et an enhancer trap dissociation transposon, GT a gene
trap dissociation transposon, MT a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. 3 denotes a sequence derived from the 3' end of the
transposon, _5 denotes a sequence derived from the 5' end of the
transposon. BESRC GARNET, ATIS project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
code: N104131.
FEATURES
source
Location/Qualifiers
1..243
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AL162873"
/note="Derived from superpool 23.41 NASC code N41088"
STS
1..243
/standard_name="SM_3.16284"
ORIGIN
Alignment Scores:
Pred. No.: 2.52e+03 Length: 243
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0

```

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Morris,D.W. and Engelhard,E.K.
TITLE Novel compositions and methods for cancer
JOURNAL Patent: WO 03008583-A 293 30-JAN-2003;
Sagres Discovery (US)
FEATURES
Location/Qualifiers
source
1. .247
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.: 2.57e+03 Length: 247
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0
US-10-030-194A-6 (1-6) x AX694666 (1-247)
QY 1 GlyTyr***ValGlu 5
DB 194 GGGTACGCGGTAGAG 180
RESULT 71
LOCUS HSB0192H9 248 bp DNA linear STS 23-MAR-1996
DEFINITION H.sapiens (D16S3075) DNA segment containing (CA) repeat; clone
AFM00192h9; single read, sequence tagged site.
ACCESSION Z53115
VERSION Z53115.1 GI:1234415
KEYWORDS STS; CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;
microsatellite marker; repeat polymorphism.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Dib,C., Faure,S., Fitzames,C., Samson,D., Drouot,N., Vignal,A.,
Millasseau,P., Marc,S., Hazan,J., Seboun,E., Lathrop,M., Gyapay,G.,
Morissette,J. and Weissenbach,J.
TITLE A comprehensive genetic map of the human genome based on 5,264
microsatellites
JOURNAL Nature 380 (6570), 152-154 (1996)
MEDLINE 96176476
PUBMED 8600387
REFERENCE
AUTHORS Weissenbach,J.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1995) Genethon, B.P. 60, 91002 Evry Cedex France.
E-mail: Jean.Weissenbach@genethon.fr
COMMENT full automatic.
FEATURES
Location/Qualifiers
source
1. .248
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/cell_line="CEPH 134702"
/clone_lib="genomic DNA"
/notes="Cloning vector is M13mp18"
ORIGIN
Alignment Scores:
Pred. No.: 2.58e+03 Length: 248
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1

```

```

Query Match: 91.67% Indels: 0
DB: 11 Gaps: 0
US-10-030-194A-6 (1-6) x HSB0192H9 (1-248)
QY 1 GlyTyr***ValGlu 5
DB 137 GGTATACAGTAGAG 151
RESULT 72
LOCUS HSB184G6R 250 bp DNA linear PRI 18-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 184g6,
reverse read cpg184g6.r11a.
ACCESSION Z57571
VERSION Z57571.1 GI:1028802
KEYWORDS CpG island; genomic MseI fragment.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
PUBMED 8012384
REFERENCE
AUTHORS Dodsworth,S.J., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
COMMENT Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES
Location/Qualifiers
source
1. .250
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="184g6"
/sex="male"
/tissue_type="blood"
/clone_lib="CGI-1"
/dev_stage="adult"
ORIGIN
Alignment Scores:
Pred. No.: 2.6e+03 Length: 250
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0
US-10-030-194A-6 (1-6) x HSB184G6R (1-250)
QY 1 GlyTyr***ValGlu 5
DB 178 GGCTACAGCGTCGAA 192
RESULT 73
LOCUS AX188844 251 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 45 from Patent WO0148209.
ACCESSION AX188844
VERSION AX188844.1 GI:15142385
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 3, 2004, 11:22:27 ; Search time 314 Seconds  
(without alignments)  
100.307 Million cell updates/sec

Title: US-10-030-194A-6

Perfect score: 24  
Sequence: 1 GYVEX 6

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/US10030194/runat\_01112004\_184846\_28945/app\_query.fasta\_1.398  
-DB=N\_Geneseq\_23Sep04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=1000 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10030194 @CGN 1 1 1063 @runat\_01112004\_184846\_28945 -NCPV=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_23Sep04:\*

1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002as: \*  
7: Geneseqn2002bs: \*  
8: Geneseqn2003as: \*  
9: Geneseqn2003bs: \*  
10: Geneseqn2003cs: \*  
11: Geneseqn2003ds: \*  
12: Geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	91.7	20	12	ADJ85253
C 2	22	91.7	21	2	AAV57637 Exon 3 of
C 3	22	91.7	21	2	AAV57637 Exon 3 of
C 4	22	91.7	25	9	ACI30025 Human mic
5	22	91.7	28	4	AA504564 Gene expr
6	22	91.7	31	8	ABV93902 Bacillus

31	10	ADK71468	Adk71468 Drug-tole
35	8	ABV93903	Abv93903 Bacillus
48	2	AAV82666	Aav82666 Target ol
60	6	ABN49456	Abn49456 Human spl
60	6	ABN40992	Abn40992 Human spl
65	6	ABN30811	Abn30811 Rat splic
75	2	AAV77649	Aav77649 Staphyloc
83	8	ABX55420	Abx55420 Bovine ES
88	1	AAAN0971	Aan0971 Synthetic
88	12	ACH88299	Ach88299 Human gen
95	2	AAT30633	Aat30633 Probe nuc
100	3	AAAC67969	Aac67969 Human alp
100	3	AAAC67972	Aac67972 Human alp
100	8	ACD81433	Acd81433 E. coli K
100	8	ACD81435	Acd81435 E. coli K
100	8	ACD81434	Acd81434 E. coli K
100	8	ABS57893	Abs57893 Human alp
100	8	ABS57890	Abs57890 Human alp
108	4	AAAF75449	Aaf75449 Codon-opt
121	12	ADK33589	Adk33589 Polynucle
124	12	ACH84646	Ach84646 Human gen
126	4	ABN80737	Abn80737 Unique ba
127	4	ABN80738	Abn80738 Unique ba
127	4	ABN80738	Abn80738 Unique ba
136	4	ABA73680	Aba73680 Human foe
136	4	AAI54119	Aai54119 Probe #22
136	4	AAK48294	Aak48294 Human bon
136	4	AAK22127	Aak22127 Human bra
136	4	ABS47995	Abs47995 Human liv
166	2	AAV79014	Aav79014 Staphyloc
169	6	ABN16095	Abn16095 Human ORF
189	4	ABA69543	Aba69543 Human foe
189	4	ABA36466	Aba36466 Probe #14
189	5	AAI09933	Aai09933 Probe #99
189	6	ABS17782	Abs17782 Human gen
193	6	RAD37663	Rad37663 Maize ATP
193	8	ACD29603	Acd29603 Maize ATP
202	8	ABZ18648	Abz18648 Group III
202	10	ACA55687	ACA55687 Rabbit si
202	12	AD155483	Ad155483 Human pol
206	3	AAI31424	Aai31424 Human sec
225	10	ADH84632	Adh84632 Enterococ
227	8	ABX51080	Abx51080 Bovine ES
228	10	ADH83521	Adh83521 Enterococ
241	8	ABX41213	Abx41213 Bovine ES
247	9	ADA01774	Ada01774 Mouse car
247	10	ADB71513	Adb71513 Mouse car
251	5	AAH81246	Aah81246 Escherich
251	8	ACA13575	Aca13575 Prokaryot
265	4	AAK89055	Aak89055 Human dig
265	5	AAAS39651	Aas39651 Genomic s
265	9	ADB32611	Adb32611 Human nov
267	4	AAK87739	Aak87739 Human dig
267	5	AAAS39374	Aas39374 cDNA enco
267	9	ADB32100	Adb32100 Human nov
269	4	AAK58037	Aak58037 Human imm
273	5	AAH20035	Aah20035 Mouse dif
273	10	ABX14653	Abx14653 Mouse dif
276	8	ABZ55211	Abz55211 Aspergill
277	3	AAAC23928	Aac23928 Human sec
281	2	AAQ60834	Aaq60834 Human bra
282	8	ABZ52598	Abz52598 Aspergill
285	3	AAAC22592	Aac22592 Human sec
286	2	AAAT32663	Aat32663 E. coli p
286	2	AAAT27833	Aat27833 E. coli p
286	5	AAH81247	Aah81247 Escherich
286	8	ACA13551	Aca13551 Prokaryot
287	4	AAAD04460	Aad04460 Exon 18 o
287	4	AAH31163	Aah31163 Human ins
287	4	AAH31163	Aah31163 Human ins
287	5	AAH50563	Aah50563 Insulin r
287	5	AAQ72718	Aaq72718 Human ins
287	9	AAAL62765	Aal62765 Human ins

c 80	22	91.7	292	2	AAT25120	Human gen	153	22	91.7	395	4	AAL136922	Aai136922	Probe #56
c 81	22	91.7	292	3	AZA46937	Human sac	154	22	91.7	395	4	ABA26909	AbA26909	Probe #53
c 82	22	91.7	294	6	AQX90943	M. capsul	155	22	91.7	395	4	AkA31011	Aak31011	Human bon
c 83	22	91.7	300	2	AQX98361	Human can	156	22	91.7	395	4	AkO5412	Ako5412	Human bra
c 84	22	91.7	305	4	AAL11634	Human bre	157	22	91.7	395	4	ABs30689	AbS30689	Human liv
c 85	22	91.7	309	6	ABN76041	Human ORF	158	22	91.7	395	6	ABN21213	AbN21213	Human ORF
c 86	22	91.7	314	2	AAT14010	LERK-6 ex	159	22	91.7	395	6	ABs05760	AbS05760	Human gen
c 87	22	91.7	314	2	AAX32766	Exon sequ	160	22	91.7	396	4	AAS24335	AA24335	Human ova
c 88	22	91.7	318	6	ABN76795	Human red	161	22	91.7	400	5	ABV07265	ABv07265	Human pro
c 89	22	91.7	319	3	AA442914	Human sec	162	22	91.7	401	10	ABZ84527	ABz84527	Toxicolog
c 90	22	91.7	322	3	AAZ49399	Human STR	163	22	91.7	402	4	ABA26539	ABa26539	Probe #50
c 91	22	91.7	322	5	AAZ07074	NCI CGAP	164	22	91.7	402	4	AkO5021	AkO5021	Human bra
c 92	22	91.7	330	2	AQ04657	Sequence	165	22	91.7	402	4	ABs30218	AbS30218	Human liv
c 93	22	91.7	330	2	AQ04594	AaQ04594	166	22	91.7	405	2	AAT90986	Aat90986	Nucleotid
c 94	22	91.7	330	2	AQ068744	CH2255 li	167	22	91.7	406	9	ACH17565	Ach17565	Human adu
c 95	22	91.7	330	3	AA339597	AaQ39977	168	22	91.7	411	8	ABZ56656	ABz56656	Aspergill
c 96	22	91.7	336	2	ABX34588	Mouse MAB	169	22	91.7	413	4	AA191251	AAi91251	Human pol
c 97	22	91.7	339	5	ABV58795	Human pro	170	22	91.7	414	11	ACH95662	Ach95662	Klebsiell
c 98	22	91.7	342	2	AAX34465	Mouse MAB	171	22	91.7	415	3	AAC55086	Aac55086	Arabidops
c 99	22	91.7	342	6	ABN15948	Human ORF	172	22	91.7	416	5	ADI74723	Adi74723	Human ova
c 100	22	91.7	342	10	ADH83522	Adh83522	173	22	91.7	417	9	ACH17756	Ach17756	Human adu
c 101	22	91.7	346	6	ABQ66873	Human DNA	174	22	91.7	418	6	ABK35303	AbK35303	Human cDN
c 102	22	91.7	346	6	ABQ66873	Human DNA	175	22	91.7	418	12	ADL12783	AdL12783	Human etc
c 103	22	91.7	348	10	ADH84913	Adh84913	176	22	91.7	420	5	AAS83338	Aas83338	DNA encod
c 104	22	91.7	348	10	ADH84913	Adh84913	177	22	91.7	420	11	ADH82864	AdH82864	Enterococ
c 105	22	91.7	348	11	ABD13891	Abd13891	178	22	91.7	422	10	ACH18776	Ach18776	Rice endo
c 106	22	91.7	353	5	AA65761	Novel hum	180	22	91.7	422	9	ABV50290	ABv50290	Human pro
c 107	22	91.7	356	12	ADQ19578	Human sof	181	22	91.7	428	5	ABV50290	ABv50290	Human pro
c 108	22	91.7	357	6	ABZ13917	Abz13917	182	22	91.7	430	8	ABK39741	AbK39741	Bovine ES
c 109	22	91.7	360	4	AAI12215	Probe #21	183	22	91.7	431	2	ABV86258	ABv86258	EST clone
c 110	22	91.7	360	4	ABA53921	Human foe	184	22	91.7	431	5	ABV48757	ABv48757	Human pro
c 111	22	91.7	360	4	AAI33567	Probe #22	185	22	91.7	431	5	ABV48757	ABv48757	Human pro
c 112	22	91.7	360	4	ABM43469	Human bre	186	22	91.7	432	11	ABD17810	ABd17810	Pseudomon
c 113	22	91.7	360	4	ABA23668	Probe #21	187	22	91.7	434	8	ABX42880	ABx42880	Bovine ES
c 114	22	91.7	360	4	AAK27636	Human bon	188	22	91.7	436	5	ADI70561	Adi70561	Human gen
c 115	22	91.7	360	4	ABS27206	Human bra	189	22	91.7	436	5	ADI70561	Adi70561	Human ova
c 116	22	91.7	360	4	ABS27206	Human liv	190	22	91.7	437	5	ADI76702	Adi76702	Human ova
c 117	22	91.7	360	5	AAI02127	Probe #21	191	22	91.7	437	5	ADI70376	Adi70376	Human ova
c 118	22	91.7	366	2	AAV58685	Aav58685	192	22	91.7	447	8	ACA36565	Aca36565	Prokaryot
c 119	22	91.7	366	2	AAV61349	Reverse D	193	22	91.7	449	4	AAI79979	Aai79979	Human pol
c 120	22	91.7	366	3	AAAO6448	Human imm	194	22	91.7	459	6	ABN66777	ABn66777	Streptoco
c 121	22	91.7	366	3	AB571353	Probe #21	195	22	91.7	460	5	ABAL13307	ABa13307	Human ner
c 122	22	91.7	366	4	AAH93564	Human pro	196	22	91.7	460	6	ABL77316	ABl77316	Human ova
c 123	22	91.7	366	4	AAH93564	Human pro	197	22	91.7	463	3	AAC35950	Aac35950	Arabidops
c 124	22	91.7	366	4	AAH02629	Prostate	198	22	91.7	464	3	AAC35950	Aac35950	Arabidops
c 125	22	91.7	366	4	AAH84878	Human pro	199	22	91.7	464	4	AAI82193	Aai82193	Human pol
c 126	22	91.7	366	5	ACA59465	Prostate	200	22	91.7	467	6	ABK63196	AbK63196	Rat sequ
c 127	22	91.7	366	5	AAI50207	Human pro	201	22	91.7	467	10	ADB51960	ABs1960	Primary r
c 128	22	91.7	366	6	ABL95028	Human 8-b	202	22	91.7	470	4	AAL16280	Aal16280	Human bre
c 129	22	91.7	366	6	AB58737	Prostate	203	22	91.7	470	10	ADB55773	ABs5773	Toxicity-
c 130	22	91.7	366	8	ACC95192	Prostate	204	22	91.7	471	3	AAC37253	Aac37253	Arabidops
c 131	22	91.7	366	10	ADB13665	Human pro	205	22	91.7	471	3	AAC37253	Aac37253	Arabidops
c 132	22	91.7	368	6	ABQ97176	Mouse ES	206	22	91.7	472	2	AAV77487	Aav77487	Staphyloc
c 133	22	91.7	368	6	ABQ97176	Mouse ES	207	22	91.7	473	10	ADD35077	Add35077	Mouse mlt
c 134	22	91.7	369	3	ADP56838	Orogenita	208	22	91.7	475	9	ACH40690	Ach40690	Human foe
c 135	22	91.7	371	8	ABZ54453	Aspergill	209	22	91.7	476	4	ABA56981	AbA56981	Human foe
c 136	22	91.7	372	2	AAQ77872	Neural th	210	22	91.7	476	4	ABA26596	ABa26596	Probe #50
c 137	22	91.7	372	2	AAQ77872	AD 2-2 hu	211	22	91.7	476	5	AAI04956	Aai04956	Probe #49
c 138	22	91.7	379	8	ABZ52159	Aspergill	212	22	91.7	479	6	ABs05281	ABs05281	Human gen
c 139	22	91.7	384	4	AAAS3970	Human ova	213	22	91.7	479	6	ABK77009	ABk77009	Bacillus
c 140	22	91.7	384	6	ABK39720	Human gen	214	22	91.7	480	1	AAAN91663	Aan91663	Light cha
c 141	22	91.7	384	8	ACA12049	Human lun	215	22	91.7	484	5	AAH82924	Aah82924	Human ova
c 142	22	91.7	384	8	ABX64662	Human gen	216	22	91.7	485	9	ACH43812	Ach43812	Human foe
c 143	22	91.7	384	8	ACA03235	Lung canc	217	22	91.7	486	5	AAH82533	Aah82533	Human ova
c 144	22	91.7	384	10	ADH47287	Human lun	218	22	91.7	486	5	ADL39960	AdL39960	Human ova
c 145	22	91.7	385	4	AAAS5089	Human can	219	22	91.7	486	10	ADF80641	Adf80641	Leukaemia
c 146	22	91.7	386	8	ABX37574	Bovine ES	220	22	91.7	486	11	ABD09912	ABd09912	Pseudomon
c 147	22	91.7	390	12	ADP91819	Cotton ex	221	22	91.7	490	6	ABI99741	Abi99741	Mouse lsc
c 148	22	91.7	392	3	AAH30092	Human col	222	22	91.7	490	9	ACH39998	Ach39998	Human foe
c 149	22	91.7	393	4	AAAS8912	Human can	223	22	91.7	492	4	AAI21302	Aai21302	Probe #11
c 150	22	91.7	394	8	ABX62823	Arabidops	224	22	91.7	492	4	ABa66385	AbA66385	Human foe
c 151	22	91.7	395	4	AAI5504	Probe #54	225	22	91.7	492	4	AAI46578	Aai46578	Probe #15
c 152	22	91.7	395	4	ABA57377	Human foe	225	22	91.7	492	4	ABA48480	ABa48480	Human bre



c 372	22	91.7	739	6	ABQ52389	Abq52389 Oligonucle	445	22	91.7	924	8	ACA26792	Aca26792 Prokaryot
373	22	91.7	739	6	ABQ52388	Abq52388 Oligonucle	446	22	91.7	927	9	ADA30699	Ada30699 DNA encod
374	22	91.7	742	4	AAH07871	Aah07871 Human cDN	c 447	22	91.7	931	12	ADE77045	Ade77045 Human cDN
c 375	22	91.7	743	3	AAZ297427	Aaz297427 Human pro	c 448	22	91.7	936	11	ABD02900	Abd02900 Pseudomon
376	22	91.7	743	4	AAK91696	Aak91696 Human cDN	449	22	91.7	942	10	ABZ77943	Abz77943 Human G p
377	22	91.7	743	4	AAK93885	Aak93885 Human cDN	450	22	91.7	942	11	ABD17661	Abd17661 Pseudomon
378	22	91.7	743	12	ADL28123	Adl28123 5' end of	451	22	91.7	945	5	AAH65340	Aah65340 C glutami
379	22	91.7	743	12	ADL28123	Adl28123 3' end of	452	22	91.7	945	8	ACA02044	Aca02044 C. Glutam
380	22	91.7	746	4	AAK93843	Aak93843 Human cDN	453	22	91.7	947	3	ACA54890	Aac54890 Arabidops
381	22	91.7	746	4	AAH08410	Aah08410 Human cDN	454	22	91.7	949	3	AAK33834	Aac33834 Arabidops
382	22	91.7	746	12	ADL30270	Adl30270 3' end of	c 455	22	91.7	957	10	ADI23940	Adi23940 Streptomy
c 383	22	91.7	747	4	AAK90974	Aak90974 Human dig	c 456	22	91.7	963	4	AAH31919	Aah31919 Human olf
c 384	22	91.7	747	5	AAK32009	Aas32009 Human liv	c 457	22	91.7	963	8	ACA47588	Aca47588 Prokaryot
c 385	22	91.7	747	6	ABN90364	Abn90364 Human liv	458	22	91.7	965	4	AAF80405	Aaf80405 Nucleotid
c 386	22	91.7	747	10	ADD22649	Add22649 Filamento	459	22	91.7	965	10	ADD31052	Add31052 Plant Yie
c 387	22	91.7	747	11	ADJ15277	Adj15277 Human liv	460	22	91.7	965	12	ADI43632	Adi43632 Plant tra
388	22	91.7	748	10	ADE57100	Ade57100 Rat gene	461	22	91.7	965	12	ADO02446	Ado02446 Thalecres
389	22	91.7	748	10	ADD47001	Add47001 Rat gene	462	22	91.7	966	6	ADO62968	Ado62968 Transcrip
390	22	91.7	748	10	ADE57096	Ade57096 Rat gene	463	22	91.7	966	8	ACA47319	Aca47319 Prokaryot
391	22	91.7	748	10	ADD47005	Add47005 Rat gene	464	22	91.7	966	8	ACA47280	Aca47280 Prokaryot
c 392	22	91.7	750	6	ABQ19897	Abq19897 Oligonucle	465	22	91.7	968	3	AAF14339	Aaf14339 Aspergill
393	22	91.7	750	6	ABQ19896	Abq19896 Oligonucle	c 466	22	91.7	968	3	AAH90076	Aah90076 Human bon
394	22	91.7	751	12	ADJ42597	Adj42597 Plant cDN	467	22	91.7	974	5	AAH90076	Aah90076 Human bon
395	22	91.7	752	2	AAK98920	Aax98920 Human val	468	22	91.7	975	4	AAS51892	Aas51892 Staphyloc
396	22	91.7	753	9	ADA30315	Ada30315 DNA encod	469	22	91.7	975	8	ACF74833	Acf74833 Staphyloc
397	22	91.7	754	2	ADR02078	Adr02078 A. gossyp	470	22	91.7	978	4	AAS54619	Aas54619 Staphyloc
c 398	22	91.7	755	3	AAA45954	Aaa45954 Human met	471	22	91.7	978	4	AAS52968	Aah52968 S. epider
c 399	22	91.7	759	6	ABQ50033	Abq50033 Oligonucle	472	22	91.7	978	4	AAH52652	Aah52652 S. epider
400	22	91.7	759	6	ABQ50032	Abq50032 Oligonucle	473	22	91.7	978	8	ACC85750	Acc85750 Predicted
c 401	22	91.7	768	11	ACH96047	Ach96047 Klebsiell	474	22	91.7	978	8	ACC85751	Acc85751 Experimen
c 402	22	91.7	769	12	ADQ22745	Adq22745 Human sof	475	22	91.7	978	8	ACA19850	Aca19850 Prokaryot
c 403	22	91.7	774	8	ACA37422	Aca37422 Prokaryot	476	22	91.7	987	8	ACA19850	Aca19850 Prokaryot
c 404	22	91.7	774	8	ACH18478	Aal18478 Human bre	c 477	22	91.7	987	5	AAS42359	Aas42359 Human cDN
c 405	22	91.7	785	12	ACH87548	Ach87548 Human gen	c 478	22	91.7	987	6	ABK16635	Abk16635 Human G-c
c 406	22	91.7	792	10	ACF69197	Acf69197 Phototrab	c 479	22	91.7	987	6	ABK16635	Abk16635 Human G-c
407	22	91.7	810	4	AAL09806	Aal09806 Human bre	c 480	22	91.7	987	6	ABK68526	Abk68526 Human DNA
c 408	22	91.7	814	5	ADL62658	Adl62658 Human ova	c 481	22	91.7	987	6	ABK68526	Abk68526 Human DNA
c 409	22	91.7	819	4	ABL30425	Abi30425 Drosophil	c 482	22	91.7	989	6	ABK37645	Abk37645 DNA encod
c 410	22	91.7	833	12	ADQ22329	Adq22329 Human sof	483	22	91.7	999	10	ABZ41826	Abz41826 Human UGT
c 411	22	91.7	834	6	ABQ37201	Abq37201 Oligonucle	c 484	22	91.7	999	10	ABZ41826	Abz41826 Human UGT
c 412	22	91.7	834	6	ABQ37200	Abq37200 Oligonucle	485	22	91.7	999	10	ABZ41826	Abz41826 Human UGT
c 413	22	91.7	834	6	ACA18963	Aca18963 Prokaryot	c 486	22	91.7	1001	3	AAH51536	Aah51536 Human UGT
c 414	22	91.7	840	3	ACA51881	Aac51881 Arabidops	c 487	22	91.7	1001	3	AAH51536	Aah51536 Human UGT
c 415	22	91.7	841	8	ABZ50923	Abz50923 Aspergill	c 488	22	91.7	1001	3	AAH51536	Aah51536 Human UGT
c 416	22	91.7	848	8	ACC95735	Acc95735 Prostate	c 489	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
c 417	22	91.7	848	10	ADB14443	Adb14443 Human pro	c 490	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
c 418	22	91.7	848	10	ADG26859	Adg26859 Human pro	c 491	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
c 419	22	91.7	849	10	ADC91913	Adc91913 E. faeciu	c 492	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
c 420	22	91.7	850	6	ABQ14348	Abq14348 Oligonucle	c 493	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
c 421	22	91.7	850	6	ABQ14349	Abq14349 Oligonucle	494	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
c 422	22	91.7	852	6	ABQ90229	Abq90229 M. capsul	495	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
c 423	22	91.7	853	4	AAH06309	Aah06309 Human cDN	496	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
c 424	22	91.7	857	4	ABL08631	Abi08631 Drosophil	c 497	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
425	22	91.7	857	4	AAH31914	Aah31914 Human olf	c 498	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
426	22	91.7	859	6	ABK75099	Abk75099 Bacillus	499	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
427	22	91.7	864	8	ADA71233	Ada71233 Rice gene	500	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
428	22	91.7	879	5	ADL45834	Adl45834 Human ova	501	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
429	22	91.7	881	12	ADP20322	Adp20322 Pinus rad	c 502	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
430	22	91.7	881	12	ADP20360	Adp20360 Pinus rad	503	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
431	22	91.7	888	10	ADF41713	Adf41713 Bacillus	504	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
432	22	91.7	894	8	AAD52801	Aad52801 Aspergill	505	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
433	22	91.7	894	12	ADF82796	Adf82796 Human sec	506	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
434	22	91.7	898	4	AAD08527	Aad08527 Human sec	507	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
435	22	91.7	899	3	ACA48201	Aac48201 Arabidops	508	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
436	22	91.7	900	4	ABL06353	Abi06353 Drosophil	509	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
437	22	91.7	901	3	ACA34943	Aca34943 Arabidops	510	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
438	22	91.7	903	6	ABN91122	Abn91122 Staphyloc	511	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
439	22	91.7	909	6	ABA94721	Abas94721 Human dru	512	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
440	22	91.7	913	5	AAS75613	Aas75613 DNA encod	513	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
441	22	91.7	915	8	ADA70304	Ada70304 Rice gene	514	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
442	22	91.7	917	3	AC76546	Aac76546 Human ORF	515	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
c 443	22	91.7	921	8	ACA46950	Aca46950 Prokaryot	516	22	91.7	1002	6	ADO35088	Ado35088 Human KCh
444	22	91.7	924	4	AAH53364	Aah53364 S. epider	517	22	91.7	1002	6	ADO35088	Ado35088 Human KCh

c 518	22	91.7	1066	12	ADJ43037	Adj43037 Plant cDN
c 519	22	91.7	1077	9	ADJ29575	Adj29575 DNA encod
c 520	22	91.7	1080	6	ABN67058	Abn67058 Streptoco
521	22	91.7	1083	8	ACA50443	ACA50443 Prokaryot
522	22	91.7	1092	8	ACA36302	ACA36302 Prokaryot
c 523	22	91.7	1094	10	ADE55511	Ade55511 Rat gene
c 524	22	91.7	1095	2	AAV18817	Aav18817 Human pap
c 525	22	91.7	1095	2	AAV13389	AAV13389 Enterococ
c 526	22	91.7	1095	6	AAU49892	AAU49892 Human pal
c 527	22	91.7	1095	6	ABS99184	ABs99184 Enterococ
528	22	91.7	1095	6	ABS54383	ABs54383 DNA encod
529	22	91.7	1095	10	ADC64521	Adc64521 HPV 16 DN
530	22	91.7	1098	4	APF75385	Apf75385 Codon-opt
531	22	91.7	1098	4	AAF24352	Aaf24352 Human pap
532	22	91.7	1098	4	AAF24353	Aaf24353 Human pap
533	22	91.7	1098	8	AAU52231	AAU52231 Mutant HP
534	22	91.7	1098	8	AAU52233	AAU52233 Human pap
535	22	91.7	1098	8	AAU52229	AAU52229 Mutant HP
536	22	91.7	1098	8	AAU52230	AAU52230 Mutant HP
537	22	91.7	1098	8	AAU52232	AAU52232 Mutant HP
538	22	91.7	1098	10	ADF09605	Adf09605 Human pap
539	22	91.7	1113	2	AXX89757	Axx89757 Probe seq
540	22	91.7	1119	10	ABZ40496	Abz40496 N. gonorr
541	22	91.7	1120	6	ABK72993	Abk72993 Bacillus
c 542	22	91.7	1122	3	AAZ53361	Aaz53361 Neisseria
543	22	91.7	1124	8	ACD13425	Acd13425 cDNA enco
c 544	22	91.7	1125	11	ABD17462	Abd17462 Pseudomon
545	22	91.7	1127	5	AA882408	AA882408 DNA encod
546	22	91.7	1127	5	AA881997	AA881997 DNA encod
547	22	91.7	1130	8	ACD13224	Acd13224 cDNA enco
548	22	91.7	1138	8	ACC43589	Acc43589 Nucleotid
549	22	91.7	1140	4	ABU12513	Abu12513 Drosophil
550	22	91.7	1140	12	AAQ02084	AAQ02084 Neisseria
551	22	91.7	1142	2	AAQ29105	AAQ29105 I3L promo
552	22	91.7	1146	3	AAF12326	Aaf12326 Aspergill
553	22	91.7	1149	5	AAH65898	Aah65898 C glutami
554	22	91.7	1149	5	AB878670	AB878670 S. ghanae
555	22	91.7	1149	11	ABD14704	Abd14704 Pseudomon
556	22	91.7	1151	12	ADO35401	Ado35401 Novel mou
557	22	91.7	1155	2	AAQ01082	AAQ01082 2-Alpha-f
558	22	91.7	1155	2	AAQ098461	AAQ098461 GDP-L-fuc
559	22	91.7	1155	2	AAQ01083	AAQ01083 2-Alpha-f
c 560	22	91.7	1158	6	AB865716	AB865716 Mouse gen
c 561	22	91.7	1172	5	AA593035	AA593035 DNA encod
562	22	91.7	1173	11	ABD09342	ABD09342 Pseudomon
563	22	91.7	1174	2	AAU12238	AAU12238 Human H-t
564	22	91.7	1174	3	AA53820	AA53820 Human H-t
565	22	91.7	1182	5	AAH65738	Aah65738 C glutami
566	22	91.7	1197	10	ABZ41444	Abz41444 N. gonorr
c 567	22	91.7	1200	2	AAU91452	AAU91452 Mycobacte
c 568	22	91.7	1200	2	AAU91516	AAU91516 Mycobacte
c 569	22	91.7	1200	2	AAV44349	AAV44349 Mycobacte
c 570	22	91.7	1200	2	AAV64457	AAV64457 M. tuberc
c 571	22	91.7	1200	2	AAZ19047	Aaz19047 M. tuberc
c 572	22	91.7	1200	2	AAZ19259	Aaz19259 M. tuberc
573	22	91.7	1211	6	AAU42355	AAU42355 Corynebac
574	22	91.7	1212	4	ABU18781	ABU18781 Drosophil
575	22	91.7	1213	3	AAZ46296	AAZ46296 Human pro
576	22	91.7	1218	3	AA953358	AA953358 B. cereus
577	22	91.7	1218	11	ACH95151	ACH95151 Klebsiell
c 578	22	91.7	1228	12	ADP46452	Adp46452 Human col
579	22	91.7	1230	2	AAU76029	AAU76029 Salmonell
580	22	91.7	1236	8	ADA69937	ADA69937 Rice gene
581	22	91.7	1239	6	ABN67444	ABn67444 Streptoco
c 582	22	91.7	1251	4	AAZ49434	Aaz49434 E. coli p
583	22	91.7	1251	4	ADU08503	ADU08503 Human sec
584	22	91.7	1252	12	ADO40238	Ado40238 Streptoco
c 585	22	91.7	1257	6	ABK24526	ABk24526 E1F-2alph
586	22	91.7	1258	4	ABU19851	ABU19851 Drosophil
c 587	22	91.7	1260	8	ABZ26079	ABz26079 Mouse DNA
c 588	22	91.7	1264	12	ADL12827	Adl12827 Human ste
589	22	91.7	1266	4	AAF71669	Aaf71669 Corynebac
590	22	91.7	1267	10	ADJ81631	Adj81631 Transloca
591	22	91.7	1269	12	ADJ39779	Adj39779 Plant cDN
592	22	91.7	1275	8	ABZ71144	Abz71144 S. muraya
c 593	22	91.7	1278	12	ADO26259	ADO26259 C35 promo
594	22	91.7	1280	11	ABD09277	ABD09277 Pseudomon
595	22	91.7	1290	2	AAx97877	Aax97877 Human sec
596	22	91.7	1290	10	ADJ45998	Adj45998 Novel hum
597	22	91.7	1290	12	ADP18812	ADP18812 Human sec
598	22	91.7	1296	11	ABD03066	ABD03066 Pseudomon
599	22	91.7	1296	11	ABD00756	ABD00756 Klebsiell
c 600	22	91.7	1296	12	ADO80770	ADO80770 Porcine e
c 601	22	91.7	1301	3	AA375538	AAc37538 Arabidops
602	22	91.7	1302	11	ABD09205	ABd09205 Pseudomon
c 603	22	91.7	1306	5	AA883341	AA883341 DNA encod
c 604	22	91.7	1307	6	ABK36134	ABK36134 cDNA sequ
605	22	91.7	1317	8	ACA31823	ACA31823 Prokaryot
c 606	22	91.7	1323	11	ABD08999	ABD08999 Pseudomon
607	22	91.7	1330	12	ADJ40258	Adj40258 Plant cDN
608	22	91.7	1332	12	ADG18079	ADg18079 Human mut
609	22	91.7	1332	12	ADG18075	ADg18075 Human dud
c 610	22	91.7	1333	12	ADP46451	ADp46451 Human col
611	22	91.7	1338	6	ABN69850	ABn69850 Streptoco
612	22	91.7	1347	5	AA881888	AA881888 DNA encod
613	22	91.7	1347	10	ADD46198	AdD46198 Rat gene
614	22	91.7	1347	10	ADE62406	Ade62406 Rat gene
615	22	91.7	1350	3	AAZ53973	AAz53973 Neisseria
616	22	91.7	1352	3	AAZ53975	AAz53975 Neisseria
617	22	91.7	1353	11	ACH96280	ACH96280 Klebsiell
c 618	22	91.7	1354	6	ABQ61100	ABQ61100 Zinc fing
619	22	91.7	1363	10	ADC85978	ADC85978 Human GPC
620	22	91.7	1365	6	ABK92252	ABk92252 Prostate
621	22	91.7	1365	11	ADN39255	ADn39255 Cancer/an
622	22	91.7	1368	2	AAV46288	AAv46288 A. foetid
c 623	22	91.7	1369	10	ADC87352	ADC87352 Human GPC
624	22	91.7	1371	3	AAA66004	AAa66004 E. coli p
625	22	91.7	1371	4	AA552403	AA552403 E. coli D
626	22	91.7	1371	8	ACA35370	ACA35370 Prokaryot
627	22	91.7	1371	8	ACA18557	ACA18557 Prokaryot
628	22	91.7	1373	8	ABZ71637	ABz71637 Breast sp
629	22	91.7	1373	12	ADF85879	ADF85879 Human bre
630	22	91.7	1374	11	ACH95011	ACH95011 Klebsiell
631	22	91.7	1375	12	ADO26480	ADO26480 Chimeric
c 632	22	91.7	1380	12	ADN99109	ADn99109 Novel hum
c 633	22	91.7	1380	12	ADO00678	ADO0678 Novel hum
c 634	22	91.7	1383	3	AAZ97477	AAz97477 Human pro
635	22	91.7	1385	10	ADJ45904	Adj45904 Novel hum
636	22	91.7	1385	10	ADC85802	ADC85802 Human GPC
c 637	22	91.7	1387	10	ACA36454	ACA36454 Prokaryot
c 638	22	91.7	1395	8	ACA36454	ACA36454 Prokaryot
c 639	22	91.7	1395	1	AAU71243	AAU71243 Sequence
640	22	91.7	1401	10	ADG76401	ADg76401 Human inc
641	22	91.7	1403	8	ACC95737	ACC95737 Prostate
642	22	91.7	1403	10	ADBI4445	ADb14445 Human pro
643	22	91.7	1403	10	ADG36861	ADg36861 Human pro
c 644	22	91.7	1404	2	AAQ13581	AAq13581 A. altoce
c 645	22	91.7	1404	2	AAQ20384	AAQ20384 ADH compl
646	22	91.7	1407	6	ABZ14582	ABz14582 Arabidops
647	22	91.7	1407	8	ACA41232	ACA41232 Prokaryot
648	22	91.7	1407	10	ABZ39937	ABz39937 N. gonorr
649	22	91.7	1420	8	AA54703	AA54703 Mouse SCR
650	22	91.7	1420	10	ADF53588	ADF53588 Haematopo
c 651	22	91.7	1426	6	ABL40152	ABL40152 16S ribos
c 652	22	91.7	1426	10	ADD12929	ADD12929 Bacterial
c 653	22	91.7	1430	2	AAU07384	AAU07384 Paracoccu
654	22	91.7	1434	8	ACA33652	ACA33652 Prokaryot
655	22	91.7	1435	6	ABQ69146	ABq69146 Listeria
656	22	91.7	1440	4	AAF25837	Aaf25837 R. marinu
c 657	22	91.7	1440	10	ADF00526	ADF00526 Bacterial
c 658	22	91.7	1451	2	AAQ81791	AAQ81791 Nucleotid
c 659	22	91.7	1452	2	AAU15842	AAU15842 DNA deriv
c 660	22	91.7	1452	2	AAV22896	AAv22896 DNA encod
c 661	22	91.7	1452	4	AAF82204	AAf82204 DNA sequ
c 662	22	91.7	1452	6	ABL40151	ABL40151 16S ribos
c 663	22	91.7	1452	10	ADD12928	ADD12928 Paracoccu

c 664	22	91.7	1452	12	ADL81773	Adl81773 P. aerugi	737	22	91.7	1659	2	AAQ13407	Aaq13407 Indoleace
c 665	22	91.7	1455	8	ACA37958	Aca37958 Prokaryot	738	22	91.7	1661	8	ACF74491	Acf74491 Staphyloc
c 666	22	91.7	1458	5	AAF26414	Aaf26414 Pseudomon	739	22	91.7	1661	12	ADO05147	Ado05147 A. thalia
c 667	22	91.7	1460	10	ADE71121	Ade71121 LOC57100	740	22	91.7	1662	4	AA554463	Aa554463 Staphyloc
c 668	22	91.7	1467	3	ACA42383	Aac42383 Arabidops	741	22	91.7	1662	4	ACA20007	Aca20007 Prokaryot
c 669	22	91.7	1467	5	AA588954	Aas88954 DNA encod	742	22	91.7	1662	6	ABN92092	Abn92092 Staphyloc
c 670	22	91.7	1467	5	AA592142	Aas92142 DNA encod	743	22	91.7	1674	8	ACA41097	Aca41097 Prokaryot
c 671	22	91.7	1468	5	AA588548	Aas88548 DNA encod	744	22	91.7	1674	8	ACA34444	Aca34444 Prokaryot
c 672	22	91.7	1468	5	AA573077	Aas73077 DNA encod	745	22	91.7	1680	5	AA515802	Aas15802 Human CDN
c 673	22	91.7	1468	5	AA588972	Aas88972 DNA encod	746	22	91.7	1680	5	AA5F58306	Aaf58306 Human CDP
c 674	22	91.7	1479	2	AAQ77874	Aaq77874 Neural th	747	22	91.7	1686	5	AA571196	Aas71196 DNA encod
c 675	22	91.7	1480	2	AAI27756	Aat27756 AD 2-2 hu	748	22	91.7	1695	3	AA49051	Aac49051 Arabidops
c 676	22	91.7	1481	4	AAI19933	Aai99333 Human exc	749	22	91.7	1698	10	ABZ41728	Abz41728 N. gonorr
c 677	22	91.7	1481	5	AAI163683	Aai63683 Human kid	750	22	91.7	1698	12	ADQ22782	Adq22782 Human sof
c 678	22	91.7	1482	6	ABZ14144	Abz14144 Arabidops	751	22	91.7	1703	10	ADE54901	Ade54901 Rat gene
c 679	22	91.7	1490	2	AA58672	Aas58672 Human org	752	22	91.7	1703	10	ADE54905	Ade54905 Rat gene
c 680	22	91.7	1497	4	AA6161015	Aaf61015 P. putida	753	22	91.7	1708	5	AAH50965	Aah50965 Lipolic ac
c 681	22	91.7	1497	9	ADA12179	Ada12179 DNA encod	754	22	91.7	1713	8	ABT19367	Abt19367 Aspergill
c 682	22	91.7	1497	10	AA5859823	Aad59823 Acinetoba	755	22	91.7	1714	4	AAH17536	Aah17536 Human CDN
c 683	22	91.7	1498	10	ABX17918	Abx17918 cDNA encod	756	22	91.7	1722	8	ABT21187	Abt21187 Aspergill
c 684	22	91.7	1500	11	ACH97053	Ach97053 Klebsiell	757	22	91.7	1725	5	AA515793	Aas15793 Human DNA
c 685	22	91.7	1507	12	ADQ19677	Adq19677 Human sof	758	22	91.7	1737	5	AAH66136	Aah66136 C Glutam
c 686	22	91.7	1512	3	AA36572	Aac36572 Arabidops	759	22	91.7	1737	10	ADL13691	Adl13691 Osteoarth
c 687	22	91.7	1514	4	AA16586	Aad16586 Human nov	760	22	91.7	1737	10	ADL13692	Adl13692 Osteoarth
c 688	22	91.7	1514	10	ADG62921	Adg62921 cDNA encod	761	22	91.7	1743	9	ACH00821	Ach00821 Human pro
c 689	22	91.7	1517	4	ABLI14229	Abli14229 Drosophil	762	22	91.7	1746	8	ACF73627	Acf73627 Staphyloc
c 690	22	91.7	1517	11	ACH96979	Ach96979 Klebsiell	763	22	91.7	1751	1	AAAN71021	Aan71021 Sequence
c 691	22	91.7	1518	11	ACH96979	Adj42999 Plant cDN	764	22	91.7	1751	1	AAAN71021	Abt20384 Aspergill
c 692	22	91.7	1521	12	ADJ42999	Adj42999 DNA encod	765	22	91.7	1752	8	ABT20384	Abt20982 Aspergill
c 693	22	91.7	1526	5	AA573553	Abz26077 Mouse DNA	766	22	91.7	1752	8	ABT20982	Adc93530 E. faeciu
c 694	22	91.7	1527	8	ABZ26077	Abz26077 Leukaemia	767	22	91.7	1752	10	AD93530	Aai58424 Human pol
c 695	22	91.7	1530	10	ADP81588	Adn95276 Human BEC	768	22	91.7	1757	4	AA158424	Adq98634 DNA encod
c 696	22	91.7	1530	11	ADN95276	Adn95276 Human CYP	769	22	91.7	1757	5	ADQ98634	Adb48394 Novel hum
c 697	22	91.7	1530	12	ADO26460	Ado26460 Human cyc	770	22	91.7	1757	9	ADQ98634	Aaz65010 Membrane-
c 698	22	91.7	1530	12	ADN033889	Adn03889 Antipspor	771	22	91.7	1759	3	AAZ65010	Aas64000 Human DNA
c 699	22	91.7	1530	12	ADQ09231	Adq09231 Human CCN	772	22	91.7	1759	5	AAAF44156	Aaf44156 Human PRO
c 700	22	91.7	1532	9	ACH03960	Ach03960 Human CDN	773	22	91.7	1759	5	AAAF44156	Aca89450 cDNA enco
c 701	22	91.7	1541	4	AA815818	Aaf81518 White lup	774	22	91.7	1759	8	ACA89450	Aca05775 Human sec
c 702	22	91.7	1561	5	AA515811	Aai58111 Human cDN	775	22	91.7	1759	8	ACA73460	Aca73460 Human sec
c 703	22	91.7	1568	6	ABA923359	Ab923359 Human glu	776	22	91.7	1759	8	ACA05775	Aca05775 Human sec
c 704	22	91.7	1569	4	AAH18251	Aah18251 Human CDN	777	22	91.7	1759	8	ACA66609	Aca66609 cDNA enco
c 705	22	91.7	1571	4	ABLI14515	Aah14515 Drosophil	778	22	91.7	1759	8	ACA64286	Aca64286 Novel hum
c 706	22	91.7	1572	3	AAAG5977	Aaa65977 E. coli p	779	22	91.7	1759	8	ACF19570	Acf19570 Human sec
c 707	22	91.7	1572	12	ADM99120	Adm99120 Bacterial	780	22	91.7	1759	8	ACD18217	Acd18217 Human sec
c 708	22	91.7	1575	5	AA571295	Aas71295 DNA encod	781	22	91.7	1759	8	ACF21858	Acf21858 Human sec
c 709	22	91.7	1580	4	AAI63862	Aai63862 Human pol	782	22	91.7	1759	8	ACF00175	Acf00175 Human sec
c 710	22	91.7	1580	12	ADM24413	Adm24413 Human PRO	783	22	91.7	1759	8	ACD25126	Acd25126 Novel hum
c 711	22	91.7	1589	6	ABS65715	Ab65715 Mouse gen	784	22	91.7	1759	8	ACA72232	Aca72232 Novel hum
c 712	22	91.7	1590	2	AAV19125	Aav19125 Nucleotid	785	22	91.7	1759	8	ACD04756	Acd04756 Human sec
c 713	22	91.7	1590	2	AA525066	Aax25066 Brassica	786	22	91.7	1759	8	ACA72232	Acd18217 Human sec
c 714	22	91.7	1590	6	ABA97357	Ab97357 Brassica	787	22	91.7	1759	8	ACD18217	Acd08224 Human sec
c 715	22	91.7	1590	12	AD136507	Adi36507 Brassica	788	22	91.7	1759	8	ACD08224	Aca89658 Novel hum
c 716	22	91.7	1593	11	ABDI7410	Abdi7410 Pseudomon	789	22	91.7	1759	8	ACA70100	Aca70100 Human sec
c 717	22	91.7	1594	3	AAA093387	Aaa093387 Human DNA	790	22	91.7	1759	8	ACD12322	Acd12322 Novel hum
c 718	22	91.7	1596	8	ACA44315	Aca44315 Prokaryot	791	22	91.7	1759	8	ACC74237	Acc74237 Human sec
c 719	22	91.7	1597	11	ADP65696	Adp65696 Human cre	792	22	91.7	1759	8	ACD15865	Acd15865 Human sec
c 720	22	91.7	1597	11	ADP65779	Adp65779 Human ear	793	22	91.7	1759	8	ACD25433	Acd25433 Novel hum
c 721	22	91.7	1597	12	ADQ18237	Adq18237 Human sof	794	22	91.7	1759	8	ACD17910	Acd17910 Human sec
c 722	22	91.7	1597	12	ADP21346	Adp21346 Gene CKMT	795	22	91.7	1759	8	ACC88197	Acc88197 Human sec
c 723	22	91.7	1598	12	ADQ23763	Adq23763 Human sof	796	22	91.7	1759	8	ACD21551	Acd21551 Human sec
c 724	22	91.7	1601	4	AAH13799	Aah13799 Human CDN	797	22	91.7	1759	8	ACD18618	Acd18618 Human sec
c 725	22	91.7	1608	8	ABZ26078	Abz26078 Mouse DNA	798	22	91.7	1759	8	ABX98228	Abx98228 Human CDN
c 726	22	91.7	1608	10	ADF00341	Adf00341 Bacterial	799	22	91.7	1759	8	ACD13979	Acd13979 Human PRO
c 727	22	91.7	1609	4	ABL24587	Ab124587 Drosophil	800	22	91.7	1759	8	ACD13979	Acd09759 Human sec
c 728	22	91.7	1609	10	ADC24204	Adc24204 Human NOV	801	22	91.7	1759	8	ACD09759	Acd09759 Human sec
c 729	22	91.7	1613	6	ABQ76417	Abq76417 S. cerevi	802	22	91.7	1759	8	ACC88504	Acc88504 Human sec
c 730	22	91.7	1618	6	ABQ39531	Abq39531 Oligonuecl	803	22	91.7	1759	8	ACD21244	Acd21244 Human sec
c 731	22	91.7	1618	6	ABQ39530	Abq39530 Oligonuecl	804	22	91.7	1759	8	ABX75616	Abx75616 Human CDN
c 732	22	91.7	1620	4	AAI59474	Aai59474 Human pol	805	22	91.7	1759	8	ABX97819	Abx97819 Human PRO
c 733	22	91.7	1632	10	AAZ77151	Aaz77151 Human pro	806	22	91.7	1759	8	ACA97295	Aca97295 Novel hum
c 734	22	91.7	1635	4	AA551954	Aas51954 Staphyloc	807	22	91.7	1759	8	ACA57758	Aca57758 Human PRO
c 735	22	91.7	1638	4	AAF29141	Aaf29141 Polynucle	808	22	91.7	1759	8	ACD14286	Acd14286 Human PRO
c 736	22	91.7	1647	8	ACA24757	Aca24757 Prokaryot	809	22	91.7	1759	8	ACC91069	Acc91069 Human sec

810	22	91.7	1759	8	ACC88811	Human sec	Acc88811	Human sec	883	22	91.7	1759	9	ACD81839	CDNA enco	Acc81839	CDNA enco
811	22	91.7	1759	8	ACD07008	Human PRO	Acc07008	Human PRO	884	22	91.7	1759	9	ACD11708	Human sec	Acc11708	Human sec
812	22	91.7	1759	8	ACA67459	Human PRO	Acc67459	Human PRO	885	22	91.7	1759	9	ACC95837	Human sec	Acc95837	Human sec
813	22	91.7	1759	8	ACC81514	Human sec	Acc81514	Human sec	886	22	91.7	1759	9	ACF16400	Human sec	Acc16400	Human sec
814	22	91.7	1759	8	ACC89118	Human sec	Acc89118	Human sec	887	22	91.7	1759	9	ACF02518	Human sec	Acc02518	Human sec
815	22	91.7	1759	8	ACC86474	Human sec	Acc86474	Human sec	888	22	91.7	1759	9	ACF02825	Human sec	Acc02825	Human sec
816	22	91.7	1759	8	ACC89732	Human sec	Acc89732	Human sec	889	22	91.7	1759	9	ACF21412	Human sec	Acc21412	Human sec
817	22	91.7	1759	8	ACC92911	Human sec	Acc92911	Human sec	890	22	91.7	1759	9	ACF10096	Human sec	Acc10096	Human sec
818	22	91.7	1759	8	ABX80745	Human sec	Abx80745	Human sec	891	22	91.7	1759	9	ACF77989	Human sec	Acc77989	Human sec
819	22	91.7	1759	8	ACA72539	Human PRO	Acc72539	Human PRO	892	22	91.7	1759	9	ACD46694	Human sec	Acc46694	Human sec
820	22	91.7	1759	8	ACA89057	Human sec	Acc89057	Human sec	893	22	91.7	1759	9	ACD49457	Human sec	Acc49457	Human sec
821	22	91.7	1759	8	ACA69793	Human sec	Acc69793	Human sec	894	22	91.7	1759	9	ACD28224	Human sec	Acc28224	Human sec
822	22	91.7	1759	8	ACA96936	Novel hum	Acc96936	Novel hum	895	22	91.7	1759	9	ACD88914	Human sec	Acc88914	Human sec
823	22	91.7	1759	8	ACA90932	Novel hum	Acc90932	Novel hum	896	22	91.7	1759	9	ACD84309	Human PRO	Acc84309	Human PRO
824	22	91.7	1759	8	ACA70714	Human sec	Acc70714	Human sec	897	22	91.7	1759	9	ACD99083	CDNA enco	Acc99083	CDNA enco
825	22	91.7	1759	8	ACA95524	Novel hum	Acc95524	Novel hum	898	22	91.7	1759	9	ADA77903	Human sec	Acc77903	Human sec
826	22	91.7	1759	8	ACD44254	CDNA enco	Acc44254	CDNA enco	899	22	91.7	1759	9	ACF48825	Human sec	Acc48825	Human sec
827	22	91.7	1759	8	ACC86167	Human sec	Acc86167	Human sec	900	22	91.7	1759	9	ACD09145	Human sec	Acc09145	Human sec
828	22	91.7	1759	8	ACC90039	Human sec	Acc90039	Human sec	901	22	91.7	1759	9	ACF11938	Human sec	Acc11938	Human sec
829	22	91.7	1759	8	ACD12647	Human sec	Acc12647	Human sec	902	22	91.7	1759	9	ACF41172	Human sec	Acc41172	Human sec
830	22	91.7	1759	8	ACF19877	Human sec	Acc19877	Human sec	903	22	91.7	1759	9	ACF15786	Human sec	Acc15786	Human sec
831	22	91.7	1759	8	ABX76821	Human PRO	Abx76821	Human PRO	904	22	91.7	1759	9	ACF16093	Human sec	Acc16093	Human sec
832	22	91.7	1759	8	ACA73153	Novel hum	Acc73153	Novel hum	905	22	91.7	1759	9	ACD31920	Human sec	Acc31920	Human sec
833	22	91.7	1759	8	ACA68696	Novel hum	Acc68696	Novel hum	906	22	91.7	1759	9	ACF18728	Human sec	Acc18728	Human sec
834	22	91.7	1759	8	ACA74540	CDNA enco	Acc74540	CDNA enco	907	22	91.7	1759	9	ACF09175	Human sec	Acc09175	Human sec
835	22	91.7	1759	8	ACA70407	Human sec	Acc70407	Human sec	908	22	91.7	1759	9	ACF78296	Human sec	Acc78296	Human sec
836	22	91.7	1759	8	ACD14593	Human PRO	Acc14593	Human PRO	909	22	91.7	1759	9	ACF51895	Human sec	Acc51895	Human sec
837	22	91.7	1759	8	ACA68265	Novel hum	Acc68265	Novel hum	910	22	91.7	1759	9	ACF26382	Human sec	Acc26382	Human sec
838	22	91.7	1759	8	ABX98730	Novel hum	Abx98730	Novel hum	911	22	91.7	1759	9	ACF24175	Human sec	Acc24175	Human sec
839	22	91.7	1759	8	ACC81207	Human sec	Acc81207	Human sec	912	22	91.7	1759	9	ACF63486	Human sec	Acc63486	Human sec
840	22	91.7	1759	8	ACA95531	Novel hum	Acc95531	Novel hum	913	22	91.7	1759	9	ACF50360	Human sec	Acc50360	Human sec
841	22	91.7	1759	8	ACD04449	Novel hum	Acc04449	Novel hum	914	22	91.7	1759	9	ACH07831	Human sec	Acc07831	Human sec
842	22	91.7	1759	8	ACC87890	Human sec	Acc87890	Human sec	915	22	91.7	1759	9	ACF13637	Human sec	Acc13637	Human sec
843	22	91.7	1759	8	ACF12552	Human sec	Acc12552	Human sec	916	22	91.7	1759	9	ACD41563	Human sec	Acc41563	Human sec
844	22	91.7	1759	8	ABX79425	Human sec	Abx79425	Human sec	917	22	91.7	1759	9	ADA337691	Human CDN	Ada337691	Human CDN
845	22	91.7	1759	8	ACA96267	Human PRO	Acc96267	Human PRO	918	22	91.7	1759	9	ACF31976	Human sec	Acc31976	Human sec
846	22	91.7	1759	8	ACA65041	Human PRO	Acc65041	Human PRO	919	22	91.7	1759	9	ACF23254	Human sec	Acc23254	Human sec
847	22	91.7	1759	8	ACA73767	Human sec	Acc73767	Human sec	920	22	91.7	1759	9	ACF39944	Human sec	Acc39944	Human sec
848	22	91.7	1759	8	ACA74179	Novel hum	Acc74179	Novel hum	921	22	91.7	1759	9	ACD45466	Human sec	Acc45466	Human sec
849	22	91.7	1759	8	ACA96574	Human PRO	Acc96574	Human PRO	922	22	91.7	1759	9	ACF53123	Human sec	Acc53123	Human sec
850	22	91.7	1759	8	ACD10680	CDNA enco	Acc10680	CDNA enco	923	22	91.7	1759	9	ACF27303	Human sec	Acc27303	Human sec
851	22	91.7	1759	8	ACC91376	Human sec	Acc91376	Human sec	924	22	91.7	1759	9	ACF45141	Human sec	Acc45141	Human sec
852	22	91.7	1759	8	ACA93446	Novel hum	Acc93446	Novel hum	925	22	91.7	1759	9	ACD29759	Human sec	Acc29759	Human sec
853	22	91.7	1759	8	ACD02711	CDNA enco	Acc02711	CDNA enco	926	22	91.7	1759	9	ACD89835	Human sec	Acc89835	Human sec
854	22	91.7	1759	8	ACC87276	Human sec	Acc87276	Human sec	927	22	91.7	1759	9	ACD84616	Human PRO	Acc84616	Human PRO
855	22	91.7	1759	8	ACC85860	Human sec	Acc85860	Human sec	928	22	91.7	1759	9	ACD98776	CDNA enco	Acc98776	CDNA enco
856	22	91.7	1759	8	ABX81128	Novel hum	Abx81128	Novel hum	929	22	91.7	1759	9	ACF77068	Human sec	Acc77068	Human sec
857	22	91.7	1759	8	ACA65348	Human PRO	Acc65348	Human PRO	930	22	91.7	1759	9	ACF76761	Human sec	Acc76761	Human sec
858	22	91.7	1759	8	ACA94165	Human sec	Acc94165	Human sec	931	22	91.7	1759	9	ACF49746	Human sec	Acc49746	Human sec
859	22	91.7	1759	8	ACA97909	Human PRO	Acc97909	Human PRO	932	22	91.7	1759	9	ACF50053	Human sec	Acc50053	Human sec
860	22	91.7	1759	8	ACA91411	Novel hum	Acc91411	Novel hum	933	22	91.7	1759	9	ADA21377	Human CDN	Ada21377	Human CDN
861	22	91.7	1759	8	ACA90625	Novel hum	Acc90625	Novel hum	934	22	91.7	1759	9	ACD09452	Human sec	Acc09452	Human sec
862	22	91.7	1759	8	ACD16172	Human sec	Acc16172	Human sec	935	22	91.7	1759	9	ACD08531	Human sec	Acc08531	Human sec
863	22	91.7	1759	8	ACD17333	Human sec	Acc17333	Human sec	936	22	91.7	1759	9	ACF12245	Human sec	Acc12245	Human sec
864	22	91.7	1759	8	ACC91990	Human sec	Acc91990	Human sec	937	22	91.7	1759	9	ACD22472	Human sec	Acc22472	Human sec
865	22	91.7	1759	8	ACA74847	CDNA enco	Acc74847	CDNA enco	938	22	91.7	1759	9	ACD94753	Human sec	Acc94753	Human sec
866	22	91.7	1759	8	ACA91718	Human PRO	Acc91718	Human PRO	939	22	91.7	1759	9	ACF15172	Human sec	Acc15172	Human sec
867	22	91.7	1759	8	ACA71362	Human sec	Acc71362	Human sec	940	22	91.7	1759	9	ACC97267	Human sec	Acc97267	Human sec
868	22	91.7	1759	8	ACC90762	Human sec	Acc90762	Human sec	941	22	91.7	1759	9	ACC92297	Human sec	Acc92297	Human sec
869	22	91.7	1759	8	ACA65772	CDNA enco	Acc65772	CDNA enco	942	22	91.7	1759	9	ACF13944	Human sec	Acc13944	Human sec
870	22	91.7	1759	8	ACA92944	Novel hum	Acc92944	Novel hum	943	22	91.7	1759	9	ACF14251	Human sec	Acc14251	Human sec
871	22	91.7	1759	8	ACA94917	CDNA enco	Acc94917	CDNA enco	944	22	91.7	1759	9	ADA10164	Human CDN	Ada10164	Human CDN
872	22	91.7	1759	8	ACD16479	Human sec	Acc16479	Human sec	945	22	91.7	1759	9	ACF09482	Human sec	Acc09482	Human sec
873	22	91.7	1759	8	ACD15558	Human sec	Acc15558	Human sec	946	22	91.7	1759	9	ACD45773	Human sec	Acc45773	Human sec
874	22	91.7	1759	8	ABX17028	Human PRO	Abx17028	Human PRO	947	22	91.7	1759	9	ACD47922	Human sec	Acc47922	Human sec
875	22	91.7	1759	8	ABX16661	Human CDN	Abx16661	Human CDN	948	22	91.7	1759	9	ACD67653	CDNA enco	Acc67653	CDNA enco
876	22	91.7	1759	9	ACA67883	Novel hum	Acc67883	Novel hum	949	22	91.7	1759	9	ACF25461	Human sec	Acc25461	Human sec
877	22	91.7	1759	9	ACA97602	Human PRO	Acc97602	Human PRO	950	22	91.7	1759	9	ACF29145	Human sec	Acc29145	Human sec
878	22	91.7	1759	9	ACA99051	Novel hum	Acc99051	Novel hum	951	22	91.7	1759	9	ACD84923	Human sec	Acc84923	Human sec
879	22	91.7	1759	9	ACC91683	Human sec	Acc91683	Human sec	952	22	91.7	1759	9	ACD84002	Human PRO	Acc84002	Human PRO
880	22	91.7	1759	9	ACD11094	Novel hum	Acc11094	Novel hum	953	22	91.7	1759	9	ACD87993	Human sec	Acc87993	Human sec
881	22	91.7	1759	9	ACD14944	Human sec	Acc14944	Human sec	954	22	91.7	1759	9	ACF30680	Human sec	Acc30680	Human sec
882	22	91.7	1759	9	ACA88332	Human sec	Acc88332	Human sec	955	22	91.7	1759	9	ACF32283	Human sec	Acc32283	Human sec

956 22 91.7 1759 9 ACH11943 cDNA enco  
957 22 91.7 1759 9 ACH12250 cDNA enco  
958 22 91.7 1759 9 ACD40642 Human sec  
959 22 91.7 1759 9 ADA17708 cDNA enco  
960 22 91.7 1759 9 ACF18114 Human sec  
961 22 91.7 1759 9 ACF08561 Human sec  
962 22 91.7 1759 9 ACF31362 Human sec  
963 22 91.7 1759 9 ACF52202 Human sec  
964 22 91.7 1759 9 ACD50071 Human sec  
965 22 91.7 1759 9 ACF38774 Human sec  
966 22 91.7 1759 9 ACF26689 Human sec  
967 22 91.7 1759 9 ACF24789 Human sec  
968 22 91.7 1759 9 ACF46369 Human sec  
969 22 91.7 1759 9 ACF27917 Human sec  
970 22 91.7 1759 9 ACD89221 Human sec  
971 22 91.7 1759 9 ACF63793 Human sec  
972 22 91.7 1759 9 ACF60433 Human sec  
973 22 91.7 1759 9 ACH12557 cDNA enco  
974 22 91.7 1759 9 ACH09980 Human sec  
975 22 91.7 1759 9 ACD03835 Human sec  
976 22 91.7 1759 9 ACD10373 Human sec  
977 22 91.7 1759 9 ACD12015 Human sec  
978 22 91.7 1759 9 ACF42400 Human sec  
979 22 91.7 1759 9 ADA27816 Human sec  
980 22 91.7 1759 9 ACF18421 Human sec  
981 22 91.7 1759 9 ACF02211 Human sec  
982 22 91.7 1759 9 ACF21719 Human sec  
983 22 91.7 1759 9 ACF10403 Human sec  
984 22 91.7 1759 9 ACF33855 Human sec  
985 22 91.7 1759 9 ACF44817 Human sec  
986 22 91.7 1759 9 ACD90449 Human sec  
987 22 91.7 1759 9 ACD91062 Human sec  
988 22 91.7 1759 9 ACF30373 Human sec  
989 22 91.7 1759 9 ACD87072 Human sec  
990 22 91.7 1759 9 ACF60126 Human sec  
991 22 91.7 1759 9 ACF46676 Human sec  
992 22 91.7 1759 9 ACF75533 Human sec  
993 22 91.7 1759 9 ADA79695 Human sec  
994 22 91.7 1759 9 ACF17193 Human sec  
995 22 91.7 1759 9 ACF22947 Human sec  
996 22 91.7 1759 9 ACF07947 Human sec  
997 22 91.7 1759 9 ACF08254 Human sec  
998 22 91.7 1759 9 ACF40558 Human sec  
999 22 91.7 1759 9 ACF53737 Human sec  
1000 22 91.7 1759 12 ADJ64560 Human PRO

ALIGNMENTS

RESULT 1  
ADJ85253  
ID ADJ85253 standard; DNA; 20 BP.  
XX AC ADJ85253;  
XX AC  
XX DT 06-MAY-2004 (first entry)  
XX DE Nucleic acid analysis-related Tag probe SeqID321.  
XX XX restriction endonuclease site; Tag gene: Poly A site;  
XX KW T7 Promoter; nucleic acid analysis; synthetic Tag gene; assay control;  
XX KW assay development; product development; product validation;  
XX KW quality control; probe; ss.  
XX XX  
XX OS Synthetic.  
XX OS Unidentified.  
XX XX  
XX FN WO2004007684-A2.  
XX XX  
XX PD 22-JAN-2004.  
XX XX  
XX PF 14-JUL-2003; 2003WO-US021990.  
XX XX

PR 12-JUL-2002; 2002US-0395530P.  
XX (AFFY-) AFFYMETRIX INC.  
PA  
XX Christians FC;  
PI  
XX WPI; 2004-122923/12.  
DR  
XX  
XX New DNA molecules made by annealing and extending overlapping 60mer  
PT oligonucleotides, useful in producing synthetic Tag genes useful as assay  
PT controls, in assay development, product development and for quality  
PT control.  
PT  
XX Disclosure; SEQ ID NO 321; 91pp; English.  
XX  
XX This invention relates to a novel DNA molecule which comprises a DNA  
CC molecule made up of the following elements in a 5' to 3' direction: a  
CC first restriction endonuclease site; a T3 promoter site; at least one Tag  
CC gene comprising at least 5 20mer Tag sequences; a Poly A site having at  
CC least 21 consecutive A residues; a second restriction endonuclease site  
CC which may be the same or different than the first restriction  
CC endonuclease site; or a T7 promoter on the opposite strand as the T3  
CC promoter. The invention may be useful in nucleic acid analysis, in  
CC particular to synthetic Tag genes useful as assay controls, in assay  
CC development, product development and validation and for quality control.  
CC The present sequence is that of a Tag oligonucleotide probe which may be  
CC used during the creation of the novel DNA molecule of the invention.  
XX  
XX Sequence 20 BP; 5 A; 6 C; 5 G; 4 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. NO.: 166 Length: 20  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 12 Gaps: 0  
US-10-030-194A-6 (1-6) x ADJ85253 (1-20)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 6 GGATACACTGTTGAG 20  
RESULT 2  
AAV57637/c  
ID AAV57637 standard; DNA; 21 BP.  
XX AC AAV57637;  
XX AC  
XX DT 27-NOV-1998 (first entry)  
XX DE  
XX Exon 3 of an ENaC subunit amplifying forward primer B-3.  
XX KW Epithelial sodium channel; ENaC; mutation; pathological condition;  
XX KW ion transport; water retention; blood pressure; metabolic acidosis;  
XX KW chronic respiratory disease; inflammation; human; PCR primer; ss.  
XX OS  
XX OS Synthetic.  
XX OS Homo sapiens.  
XX XX  
XX FN WO9840516-A1.  
XX XX  
XX PD 17-SEP-1998.  
XX XX  
XX PF 11-MAR-1998; 98WO-US004681.  
XX XX  
XX PR 11-MAR-1997; 97US-0040171P.  
XX XX (UYUA ) UNIV YALE.  
XX PI Lifton RP, Chang SS, Rossier BC;  
XX XX



DR WPI; 1998-506740/43.

PT Determination of presence of mutation conferring pathological condition

PT mediated by altered ion transport - comprises analysing sample for

PT presence of mutation of potassium ion channel gene, ENaC, or in its

PT encoded protein.

XX

PS Example 1; Page 38; 56pp; English.

XX

CC Sequences shown in AAV57601 to AAV57686 represent primers used for the

CC PCR amplification of the exons of the different subunits of the human

CC epithelial sodium channel (ENaC) gene. This is used in the method of the

CC invention of determining the presence or absence of a mutation conferring

CC a pathological condition mediated by altered ion transport. The method

CC comprises analysing a nucleic acid sample, or protein sample, for the

CC presence of a mutation in the ENaC gene, or in its encoded protein. A

CC vector containing a nucleic acid encoding a human altered variant of the

CC ENaC protein can be used to transform host cells to produce an altered

CC variant of an ENaC protein. The protein can be used to identify agents

CC that effect ion transport. The agonists can be used to treat pathological

CC conditions resulting from abnormal ion transport, such as water

CC retention, increased blood pressure, chronic respiratory and metabolic

CC acidosis and inflammation

XX

SQ Sequence 21 BP; 3 A; 8 C; 5 G; 5 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	175	Length:	21
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	2	Gaps:	0

US-10-030-194A-6 (1-6) x AAV57637 (1-21)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 19 GGCTACACTGTGGAG 5

RESULT 3

AAV57639/c

ID AAV57639 standard; DNA; 21 BP.

XX

AC AAV57639;

XX

XX 27-NOV-1998 (first entry)

XX

DE Exon 3 of an ENaC subunit amplifying forward primer B-4.

XX

XX Epithelial sodium channel; ENaC; mutation; pathological condition;

XX ion transport; water retention; blood pressure; metabolic acidosis;

XX chronic respiratory disease; inflammation; human; PCR primer; ss.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO9840516-A1.

XX

PD 17-SEP-1998.

XX

PF 11-MAR-1998; 98WO-US004681.

XX

PR 11-MAR-1997; 97US-0040171P.

XX

PA (UYIA ) UNIV YALE.

XX

PI Lifton RP, Chang SS, Rossier BC;

XX

DR WPI; 1998-506740/43.

XX

PT Determination of presence of mutation conferring pathological condition

PT mediated by altered ion transport - comprises analysing sample for

PT presence of mutation of potassium ion channel gene, ENaC, or in its

PT encoded protein.

XX

PT presence of mutation of potassium ion channel gene, ENaC, or in its

PT encoded protein.

XX

PS Example 1; Page 38; 56pp; English.

XX

CC Sequences shown in AAV57601 to AAV57686 represent primers used for the

CC PCR amplification of the exons of the different subunits of the human

CC epithelial sodium channel (ENaC) gene. This is used in the method of the

CC invention of determining the presence or absence of a mutation conferring

CC a pathological condition mediated by altered ion transport. The method

CC comprises analysing a nucleic acid sample, or protein sample, for the

CC presence of a mutation in the ENaC gene, or in its encoded protein. A

CC vector containing a nucleic acid encoding a human altered variant of the

CC ENaC protein can be used to transform host cells to produce an altered

CC variant of an ENaC protein. The protein can be used to identify agents

CC that effect ion transport. The agonists can be used to treat pathological

CC conditions resulting from abnormal ion transport, such as water

CC retention, increased blood pressure, chronic respiratory and metabolic

CC acidosis and inflammation

XX

SQ Sequence 21 BP; 3 A; 8 C; 5 G; 5 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	175	Length:	21
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	2	Gaps:	0

US-10-030-194A-6 (1-6) x AAV57639 (1-21)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 19 GGCTACACTGTGGAG 5

RESULT 4

ACI30025/c

ID ACI30025 standard; DNA; 25 BP.

XX

AC ACI30025;

XX

DT 13-OCT-2003 (first entry)

XX

DE Human microarray DNA oligonucleotide SEQ ID NO 30016.

XX

XX EST; ss; probe; expressed sequence tag; microarray; gene expression;

XX genetic variation; biallelic marker; polymorphism; human;

XX cross-species comparison.

XX

OS Homo sapiens.

XX

PN US2003104410-A1.

XX

PD 05-JUN-2003.

XX

PF 15-MAR-2002; 2002US-00098263.

XX

PR 16-MAR-2001; 2001US-0276759P.

XX

PA (AFFY-) AFFYMETRIX INC.

XX

PI Mittmann MP;

XX

DR WPI; 2003-567953/53.

XX

PT New array of nucleic acid probes, useful for in situ hybridization, in

PT Southern, Northern or dot-blot hybridization to identify or detect the

PT sequence or specific mutations of any gene.

XX

PS Claim 1; SEQ ID NO 30016; 9pp; English.

XX

CC The invention discloses a microarray comprising a plurality of nucleic

acid probes including one of 2,018,500 fully defined sequences, or its perfect match, antisense match or antisense mismatch. Also disclosed is a method of gene expression analysis. The array is used in monitoring gene expression levels by hybridisation to a DNA library, in analysis of genetic variation or in hybridisation of tag-labelled compounds. The nucleic acid probes are specifically designed for analysis of at least one target sequence. The method of analysis comprises hybridising at least one or more nucleic acids to at least two or more nucleic acid probes and detecting the hybridisation. The nucleic acid probes are attached to a solid support. The analysis comprises monitoring gene expression levels, identifying biallelic markers or polymorphisms, or family members of a gene and a cross-species comparison. Each of the nucleic acids further comprises a tag sequence. The array of nucleic acid probes is useful in situ hybridisation, in Southern, Northern or dot-blot hybridisation to identify or detect the sequence or specific mutations of any gene, in mapping the 5' termini of mRNA molecules by primer extensions or in screening cDNA or genomic libraries or subclones for additional subclones containing segments of DNA that have been isolated and previously sequenced. The sequence presented is one of the nucleic acid probes incorporated in the microarray. Note: The sequence data for this patent can also be obtained in electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)

Sequence 25 BP; 8 A; 9 C; 5 G; 3 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 211 Length: 25 Score: 22.00 Matches: 4 Percent Similarity: 80.00% Conservative: 0 Best Local Similarity: 80.00% Mismatches: 1 Query Match: 91.67% Indels: 0 DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x AC130025 (1-25)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 22 GGGTACACTGTTGAG 8

RESULT 5

AAS04564 ID AAS04564 standard; cDNA; 28 BP.

XX AAS04564;

DT 07-SEP-2001 (first entry)

XX Gene expression profile sequence #64.

XX Gene expression profile; hypersensitivity; DNA microarray; KW liver toxicity; hepatitis; tumour formation; immunosuppression; KW renal toxicity; glomerulitis; neurotoxicity; leukaemia; dementia; KW peripheral neuropathy; hypertension; hypotension; myelosuppression; KW retinopathy; inflammation; sensitisation; ss.

XX Homo sapiens.

XX WO200132928-A2.

XX 10-MAY-2001.

XX 03-NOV-2000; 2000WO-US030474.

XX 05-NOV-1999; 99US-0165398P.

XX 11-APR-2000; 2000US-0196571P.

XX (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY.

XX Farr S;

XX WPI; 2001-328806/34.

PT Identifying hypersensitivity in a subject by obtaining a gene expression

PT profile of hypersensitivity associated genes and detecting a predetermined pattern of gene expression of hypersensitivity associated genes.

XX Claim 24; Page 150; 222pp; English.

XX The sequence represents a cDNA from a gene associated with hypersensitivity to an agent, the sequence was detected in a sample by use of a DNA microarray containing genes from a gene expression profile of a DNA microarray associated with hypersensitivity to an agent. The invention relates to methods of obtaining a gene expression profile of genes associated with hypersensitivity to an agent involving comparing the gene expression profile of cells treated with the agent with the gene expression profile of cells not treated with the agent, and determining the genes that have altered expression due to exposure to the agent. Hypersensitivity in a subject can then be detected by comparing the gene expression profile of the subject with that associated with the hypersensitivity, usually by hybridisation of a sample of mRNA or cDNA from the subject to a DNA microarray containing genes from the hypersensitivity profile. The genes in the profiles are associated with liver toxicity (e.g. hepatitis), tumour formation, immunosuppression, renal toxicity (e.g. glomerulitis), neurotoxicity, leukaemia, dementia, peripheral neuropathy, hyper/hypotension, myelosuppression, retinopathy, inflammation, and sensitisation

SQ Sequence 28 BP; 5 A; 5 C; 13 G; 5 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 239 Length: 28 Score: 22.00 Matches: 4 Percent Similarity: 80.00% Conservative: 0 Best Local Similarity: 80.00% Mismatches: 1 Query Match: 91.67% Indels: 0 DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x AAS04564 (1-28)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 10 GGATATAGCGTCGAA 24

RESULT 5

ABV93902 ID ABV93902 standard; DNA; 31 BP.

XX ABV93902;

XX 08-JAN-2003 (first entry)

XX Bacillus thuringiensis toxin Cry oligonucleotide #112 SEQ ID NO:153.

XX Bacillus thuringiensis; insecticide; toxin; Cry; pepsin cleavage site; KW pepsin; PCS; ss.

XX Bacillus thuringiensis.

XX Synthetic.

XX FR2822157-A1.

XX 20-SEP-2002.

XX 19-MAR-2001; 2001FR-00003691.

XX 19-MAR-2001; 2001FR-00003691.

XX (AVET ) AVENTIS CROPS SCIENCE SA.

XX Freyssinet G, Rang C, Frutos R;

XX WPI; 2003-002439/01.

XX New modified Cry protein, useful as insecticide, comprises at least one additional pepsin cleavage site to reduce persistence in mammalian gut.

XX Example 5; Page 131; 134pp; French.

XX The present invention describes a modified Cry protein (I) that is

CC sensitive to pepsin and comprises at least one additional pepsin cleavage

CC site (PCS). Also described: (a) increasing pepsin sensitivity of Cry

CC proteins by incorporating at least one extra PCS; (b) polynucleotides

CC (II) that encode (I); (c) chimeric genes (CG) that contain a promoter,

CC (II) and terminator; (d) expression or transformation vector (III) that

CC contains CG; (e) host organism (IV) transformed with (III), also, where

CC the organism is a plant, its parts and seeds; (f) production of (I) by

CC growing (IV); and (g) mono- or polyclonal antibodies (Ab) directed

CC against (I). (I) has insecticide activity. (I) can be used as

CC insecticides, particularly where expressed in transgenic plants. (I) are

CC sensitive to enzymes in the digestive tract of mammals, so do not persist

CC in the tract (lack of persistence is required by regulatory authorities

CC for use, in foods, of seeds containing Cry proteins). Extra PCS do not

CC increase degradation in the digestive tract of insects, so have no effect

CC on insecticidal activity. ABV93450 to ABV93909 and ABP67997 to ABP68308

CC represent sequences used in the exemplification of the present invention

XX SQ Sequence 31 BP; 9 A; 4 C; 9 G; 9 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	267	Length:	31
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	8	Gaps:	0

US-10-030-194A-6 (1-6) x ABV93902 (1-31)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 15 GGTACTCCGTGGAG 29

RESULT 7

ADK71468/c

ID ADK71468 standard; DNA; 31 BP.

XX AC ADK71468;

XX DT 06-MAY-2004 (first entry)

XX DE Drug-tolerant gene related PCR primer.

XX KW detection; drug-tolerant gene; gene chip; probe; PCR; amplification;

XX OS hybridisation; primer; ss.

XX Synthetic.

XX CN1396271-A.

XX PD 12-FEB-2003.

XX PF 13-JUL-2001; 2001CN-00120441.

XX PR 13-JUL-2001; 2001CN-00120441.

XX PA (SANK-) SANXIONG HI TECH DEV CO LTD BEIJING.

XX PI Liu Y, Wang H, Li L;

XX WPI; 2003-442250/42.

XX PT Detection to drug tolerant gene by gene chip technique.

XX PS Claim 4; Page 23; 32pp; Chinese.

XX The present invention describes a process for detecting a drug-tolerant

CC gene with a gene chip technique. The method comprises fixing the DNA

CC sequence of an oligonucleotide probe to a carrier of a gene chip, the DNA

CC sequence of an elongation primer for the PCR linear amplification of a

CC target drug-tolerant gene, PCR amplification, and hybridisation of the

CC PCR resultant with a probe on the chip. The present sequence represents a

CC PCR primer which is used in the exemplification of the present invention.

XX SQ Sequence 31 BP; 5 A; 9 C; 11 G; 6 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	267	Length:	31
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	10	Gaps:	0

US-10-030-194A-6 (1-6) x ADK71468 (1-31)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 31 GGCTATCGGTCGAG 17

RESULT 8

ABV93903

ID ABV93903 standard; DNA; 35 BP.

XX AC ABV93903;

XX DT 08-JAN-2003 (first entry)

XX DE Bacillus thuringiensis toxin Cry oligonucleotide #113 SEQ ID NO:154.

XX KW Bacillus thuringiensis; insecticide; toxin; Cry; pepsin cleavage site;

XX OS pepsin; PCS; ss.

XX OS Bacillus thuringiensis.

XX Synthetic.

XX FR2822157-A1.

XX PD 20-SEP-2002.

XX PF 19-MAR-2001; 2001FR-00003691.

XX PR 19-MAR-2001; 2001FR-00003691.

XX PA (AVET) AVENTIS CROPS SCIENCE SA.

XX PI Freyssinet G, Rang C, Frutos R;

XX WPI; 2003-002439/01.

XX New modified Cry protein, useful as insecticide, comprises at least one

XX additional pepsin cleavage site to reduce persistence in mammalian gut.

XX Example 5; Page 131; 134pp; French.

XX The present invention describes a modified Cry protein (I) that is

CC sensitive to pepsin and comprises at least one additional pepsin cleavage

CC site (PCS). Also described: (a) increasing pepsin sensitivity of Cry

CC proteins by incorporating at least one extra PCS; (b) polynucleotides

CC (II) that encode (I); (c) chimeric genes (CG) that contain a promoter,

CC (II) and terminator; (d) expression or transformation vector (III) that

CC contains CG; (e) host organism (IV) transformed with (III), also, where

CC the organism is a plant, its parts and seeds; (f) production of (I) by

CC growing (IV); and (g) mono- or polyclonal antibodies (Ab) directed

CC against (I). (I) has insecticide activity. (I) can be used as

CC insecticides, particularly where expressed in transgenic plants. (I) are

CC sensitive to enzymes in the digestive tract of mammals, so do not persist

CC in the tract (lack of persistence is required by regulatory authorities

CC for use, in foods, of seeds containing Cry proteins). Extra PCS do not

CC increase degradation in the digestive tract of insects, so have no effect

CC on insecticidal activity. ABV93450 to ABV93909 and ABP67997 to ABP68308

CC represent sequences used in the exemplification of the present invention

XX  
SQ  
Sequence 35 BP: 8 A; 5 C; 12 G; 10 T; 0 U; 0 Other;

Alignment Scores:		
Prod. No.:	104	Length:
Score:	22.00	Matches:
Percent Similarity:	80.00%	Conservative:
Best Local Similarity:	80.00%	Mismatches:
Query Match:	91.67%	Indels:
DR:	8	Gaps:
		35

US-10-030-194A-6 (1-6) x ABV93903 (1-35)

Qy            1 GlyTyr\*\*\*ValGlu 5  
               |||||  
nb            5 GGGTACTCCGTGGAG 19

RESULT 9  
AAV82666/c  
IN AAV82666 standard: DNA: 48 BP.

AA  
AC  
AAV82666;

DT 25-FEB-1999 (first entry)

Target oligonucleotide JN6.

Detector oligonucleotide; hairpin structure; fluorescence; fluorescence

XX  
XX

3 XX  
by:

**FN XX**

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841

XX  
FF  
20-1544 12201

PR 30-1541-1557, XX

FA (BECI / SECTION DISCUSSION) 3  
XX

XX  
77

UR  
YY  
WEI; 1333-001100/0011

new detector. Oligo-nucleotides having 3' terminal phosphate groups are useful for PT dyes - where dyes become fluorescent when region is unpaired, useful for detection and amplification of target nucleic acid.

XX  
PC  
Example 2: Page 12: 20pp: English:

Target oligonucleotides AAV26565-68 were designed to hybridise to the detector oligonucleotide of the invention. The detector oligonucleotide comprises a single-stranded target-binding region and an intramolecularly base-paired secondary structure linked to two dyes (donor and acceptor fluorophores). In the secondary structure, fluorescence of the donor is quenched, but when it is linearised or unfolded a change in some fluorescence parameter becomes detectable. A target nucleic acid is detected by hybridizing it to a detector oligonucleotide in which the secondary structure is 5' to a target binding region, primer extension to produce a complementary strand using the secondary structure as template, resulting in linearization or unfolding of it and detecting a change in fluorescence. The detector oligonucleotide are used to detect (by primer extension and hybridization) and amplify (as signal primer) target sequences, e.g. for detecting frameshift mutations

Sequence 48 BP; 15 A; 8 C; 10 G; 15 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	428
Score:	22.00
Percent Similarity:	80.00%
Length:	48
Matches:	4
Conservative:	0

```

Best Local Similarity: 80.00%
Query Match: 91.67%
DB: 2
Mismatches: 1
Indels: 0
Gaps: 0

```

UUS-10-030-194A-6 (1-6) x AAV82666 (1-48)

Qy 1 GlyTyr\*\*\*ValGlu 5  
pb 44 GGTACTCAGTAGAG 3

## RESULT 10

ABN49456/C

AA  
AC  
AC ABN49456:

15-III, -20

Human spliced transcript d

Human. mouse: rat: splice transcript; detection; RNA transcript;   
 kw

kw splice variant; transcriptptome; transcriptome  
yy

OS Homo sapiens.

PN WO200210449-A2.

PD 07-FEB-2002.

PF 20-JUL-2001; 2001WG-1B001903.

PR 28-JUL-2000; 2000US-0221607F.

XXI — CONVENTION TWO

XX

XX  
XX

XX

PT selectively hybridize to mRNAs transcribed from a transcription unit  
PT genome, useful for detecting tissue-, pathology-, and developmental-  
PT specific genes.

Example 1: SEQ ID NO 22204; 47pp; English.

The present invention describes oligonucleotide libraries for detecting messenger RNAs that populate a (sub-)transcriptome, where the (sub-)transcriptome comprises messenger RNAs transcribed from multiple transcription units that populate a genome. The library comprises several oligonucleotides, each capable of hybridising selectively to a set of messenger RNAs transcribed from a given transcription unit of the genome, which encodes one or more messenger RNA splice variants. The oligonucleotide libraries are useful for detecting mRNAs from a biological sample, in expression profiling studies, in qualitatively or quantitatively characterising the corresponding transcriptome, and in detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini-libraries to detect transcripts of a sub-transcriptome under a particular biological or pathological state, and so allowing the detection of tissue - and pathology-specific genes such as those genes only expressed in specific tissue under a specific pathological condition; to detect developmental specific genes; and to detect RNA transcripts and splice variants of a transcriptome of a patient suffering from a particular disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from rats, humans and mice, which are used in the exemplification of the present invention. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

Sequence 60 BP; 18 A; 18 C; 10 G; 14 T; 0 U; 0 Other;

Alignment Scores: 546 Length: 60  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservativity: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB:

US-10-030-194A-6 (1-6) x ABN49456 (1-60)

Oy 1 GlyTyr\*\*\*ValGlu 5  
||||| |||||  
Db 38 GGTATTGAGTTGAG 24

RESULT 11  
ABN40992

ID ABN40992 standard; DNA; 60 BP.

XX AC ABN40992;

XX DT 15-JUL-2002 (first entry)

XX DE Human spliced transcript detection oligonucleotide SEQ ID NO:13740.

XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;

XX KW splice variant; transcriptome; oligonucleotide library; ss.

XX OS Homo sapiens.

XX PN WO200210449-A2.

XX PD 07-FEB-2002.

XX PF 20-JUL-2001; 2001WO-IB001903.

XX PR 28-JUL-2000; 2000US-0221607P.

XX PR 02-MAY-2001; 2001US-0287724P.

XX PA (COMP-) COMPUGEN INC.

XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX DR WPI; 2002-257383/30.

XX PT New oligonucleotide libraries comprising oligonucleotides which selectively hybridize to mRNAs transcribed from a transcription unit of a genome, useful for detecting tissue-, pathology-, and developmental-specific genes.

XX PS Example 1; SEQ ID NO 13740; 47pp; English.

XX CC The present invention describes oligonucleotide libraries for detecting messenger RNAs that populate a (sub-)transcriptome, where the (sub-)transcriptome comprises messenger RNAs transcribed from multiple transcription units that populate a genome. The library comprises several oligonucleotides, each capable of hybridising selectively to a set of messenger RNAs transcribed from a given transcription unit of the genome, which encodes one or more messenger RNA splice variants. The oligonucleotide libraries are useful for detecting mRNAs from a biological sample, in expression profiling studies, in qualitatively or quantitatively characterising the corresponding transcriptome, and in detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini libraries to detect transcripts of a sub-transcriptome under a particular biological or pathological state, and so allowing the detection of tissue - and pathology-specific genes such as those genes only expressed in specific tissue under a specific pathological condition; to detect developmental specific genes; and to detect RNA transcripts and splice variants of a transcriptome of a patient suffering from a particular disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from rats, humans and mice, which are used in the exemplification of the present invention. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 60 BP; 15 A; 7 C; 22 G; 16 T; 0 U; 0 Other;

Alignment Scores: 546 Length: 60  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservativity: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB:

US-10-030-194A-6 (1-6) x ABN40992 (1-60)

Oy 1 GlyTyr\*\*\*ValGlu 5  
||||| |||||  
Db 46 GGTATTGAGTTGAG 60

RESULT 12

ABN30811

ID ABN30811 standard; DNA; 65 BP.

XX AC ABN30811;

XX DT 15-JUL-2002 (first entry)

XX DE Rat spliced transcript detection oligonucleotide SEQ ID NO:3559.

XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;

XX KW splice variant; transcriptome; oligonucleotide library; ss.

XX OS Rattus norvegicus.

XX PN WO200210449-A2.

XX PD 07-FEB-2002.

XX PF 20-JUL-2001; 2001WO-IB001903.

XX PR 28-JUL-2000; 2000US-0221607P.

XX PR 02-MAY-2001; 2001US-0287724P.

XX PA (COMP-) COMPUGEN INC.

XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX DR WPI; 2002-257383/30.

XX PT New oligonucleotide libraries comprising oligonucleotides which selectively hybridize to mRNAs transcribed from a transcription unit of a genome, useful for detecting tissue-, pathology-, and developmental-specific genes.

XX PS Example 1; SEQ ID NO 3559; 47pp; English.

XX CC The present invention describes oligonucleotide libraries for detecting messenger RNAs that populate a (sub-)transcriptome, where the (sub-)transcriptome comprises messenger RNAs transcribed from multiple transcription units that populate a genome. The library comprises several oligonucleotides, each capable of hybridising selectively to a set of messenger RNAs transcribed from a given transcription unit of the genome, which encodes one or more messenger RNA splice variants. The oligonucleotide libraries are useful for detecting mRNAs from a biological sample, in expression profiling studies, in qualitatively or quantitatively characterising the corresponding transcriptome, and in detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini libraries to detect transcripts of a sub-transcriptome under a particular biological or pathological state, and so allowing the detection of tissue - and pathology-specific genes such as those genes only expressed in specific tissue under a specific pathological condition; to detect developmental specific genes; and to detect RNA transcripts and splice variants of a transcriptome of a patient suffering from a particular

CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from  
 CC rats, humans and mice, which are used in the exemplification of the  
 CC present invention. N.B. The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 65 BP; 14 A; 15 C; 22 G; 14 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 595 Length: 65  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x ABN30811 (1-65)

QY 1 GlyTyr\*\*\*ValGlu 5  
 DB 40 GGGTACAGTGTGAA 54

RESULT 13  
 AAV77649/C  
 ID AAV77649 standard; DNA; 75 BP.

XX AC AAV77649;  
 XX DT 16-MAR-1999 (first entry)  
 XX DE Staphylococcus aureus contig SEQ ID #3338.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;  
 XX cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
 KW skin infection; surgical wound infection; scalded skin syndrome;  
 KW toxic shock syndrome; ds.

XX Staphylococcus aureus.

OS EF786519-A2.

PN 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.

XX 05-JAN-1996; 96US-0009861P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -  
 PT stored on computer readable medium and used in the production of anti-  
 PT S.aureus vaccines.

XX Claim 1; Page 2524; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
 CC of the invention. The DNA sequences are recorded on a computer readable  
 CC medium, preferably selected from a floppy or hard disk, random access  
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
 CC the S.aureus DNA sequences allows putative functions to be assigned so  
 CC that protein-encoding or regulatory regions of commercial, therapeutic or  
 CC industrial importance can be obtained. Specifically, sequences which are  
 CC likely to encode antigens have been identified and these polypeptides can  
 CC be used in a vaccine composition against S.aureus infection. The  
 CC polypeptides can also be used in a kit for the immunodetection of  
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,  
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used

CC for recombinant production of the polypeptides. The new DNA sequences  
 CC (and their fragments) are useful as primers or probes for isolating  
 CC homologues of any of the S.aureus DNA sequences contained on the computer  
 CC readable medium

XX SQ Sequence 75 BP; 22 A; 14 C; 12 G; 27 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 695 Length: 75  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x AAV77649 (1-75)

QY 1 GlyTyr\*\*\*ValGlu 5  
 DB 23 GGCTATACGGTAGAA 9

RESULT 14  
 ABX55420/C  
 ID ABX55420 standard; cDNA; 83 BP.

XX AC ABX55420;  
 XX DT 26-FEB-2003 (first entry)

XX Bovine EST associated with lactation/muscle/fat deposition #5349.

XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
 KW muscle deposition; fat deposition; genome mapping; gene identification;  
 KW gene analysis; cattle breeding.

XX Bos Taurus.

XX US2002137160-A1.

XX 26-SEP-2002.

XX 26-OCT-2001; 2001US-00983965.

XX 17-DEC-1998; 98US-0113678P.

XX 15-DEC-1999; 99US-00465231.

XX (BYAT/) BYATT J C.

XX (MATH/) MATHIALAGAN N.

XX (TAON/) TAO N.

XX (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-102386/09.

XX Purified nucleic acid molecules, useful for genome mapping, gene  
 PT identification and analysis, cattle breeding or preparation of constructs  
 PT for cattle gene expression and genetically improved cattle.

XX Claim 2; SEQ ID NO 5349; 38pp; English.

XX The invention relates to a purified nucleic acid molecule associated with  
 CC lactation or muscle and fat deposition (designated LMFD), derived from  
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second  
 CC nucleic acid molecule comprising any of 5912 nucleotide sequences,  
 CC appearing as ABX50072-ABX5983, or complements of them. Also included are  
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic  
 CC acid linked to a promoter and a 3' non-translated sequence that  
 CC functions in the cell to cause termination of transcription and addition  
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
 CC (2) determining a level or pattern of a molecule in a bovine cell or  
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any  
 CC of the 5912 nucleic acid sequences or its complement or fragment) with a

CC complementary nucleic acid molecule obtained from the bovine cell or  
 CC tissue, where hybridisation between the marker nucleic acid and the  
 CC complementary nucleic acid permits the detection of the molecule; and (b)  
 CC detecting the level or pattern of the complementary nucleic acid, where  
 CC the detection of the complementary nucleic acid is predictive of the  
 CC level or pattern of the molecule. The LMFD nucleic acid is used for  
 CC determining a level or pattern of a molecule in a bovine cell or tissue.  
 CC It is useful for genome mapping, gene identification and analysis, cattle  
 CC breeding, preparation of constructs for use in cattle gene expression, or  
 CC for genetically improving cattle. The present sequence is one of the 5912  
 CC bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present  
 CC sequence was not shown in the specification but was obtained in  
 CC electronic format from the USPTO web site:  
 CC seqdata.uspto.gov/sequence.html?DocID=20020137160  
 XX

SQ Sequence 83 BP; 33 A; 19 C; 20 G; 11 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 775 Length: 83  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x ABX55420 (1-83)

QY 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 DB 72 GGTATTTCGTTGAG 58

# RESULT 15

AA80971  
 ID AAN80971 standard; DNA; 88 BP.

XX AC AAN80971;

XX 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 09-JAN-2003 (revised)

DT 12-SEP-1990 (first entry)

XX Synthetic XhoI/EcoRI fragment encoding egg white lysozyme.

XX Egg white lysozyme; hepatitis B virus (HBV) surface antigen (HBSAg);  
 KW Saccharomyces cerevisiae AH22R/pgUD LP39-Rct;  
 KW Saccharomyces cerevisiae L11P39-Rct; Saccharomyces cerevisiae LP31-Rct;  
 KW protein secretion; db.

OS Gallus gallus.

Key	Location/Qualifiers
FT misc_feature	1..13
FT	/*tag= d
FT	/note= "oligo 1"
FT misc_feature	1..4
FT	/*tag= a
FT	/note= "sticky end"
FT misc_feature	complement(5..17)
FT	/*tag= e
FT	/note= "oligo 2"
FT misc_feature	14..38
FT	/*tag= f
FT	/note= "oligo 3"
FT sig_peptide	16..69
FT	/*tag= b
FT misc_feature	complement(18..43)
FT	/*tag= g
FT	/note= "oligo 4"
FT misc_feature	39..63
FT	/*tag= h
FT	/note= "oligo 5"
FT misc_feature	complement(44..67)

FT	misc_feature	/*tag= i	/note= "Oligo 6"
FT		64..84	
FT		/*tag= j	
FT		/note= "Oligo 17"	
FT		complement(68..88)	
FT		/*tag= k	
FT		/note= "Oligo 18"	
FT		complement(85..88)	
FT		/*tag= c	
FT		/note= "sticky end"	

PN EP288198-A.

XX 26-OCT-1988.

XX 13-APR-1988; 88EP-00303297.

XX 20-APR-1987; 87JP-00098265.

PR 12-OCT-1987; 87JP-00256885.

PR 18-APR-1988; 88JP-00095335.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujisawa Y, Imai S, Miyazaki T;

XX WPI; 1988-301233/43.

DR P-PSDB; AAP80415.

XX Peptide(s) having hepatitis B surface antigenicity - used as vaccine for prevention of hepatitis B virus infection and in diagnostic kits.

XX Example; Fig 1(a); 31pp; English.

XX The signal peptide of egg white lysozyme was utilized as the signal sequence for the secretion of env protein hepatitis B virus (HBV) surface antigen (HBSAg) into a medium. A synthetic nucleotide sequence having an XhoI site at the 5'-terminal and an EcoRI site at the 3'-terminal was used, making reference to the known amino acid sequence [Jung, A. et al. P.N.A.S. 77, 5759 (1980)]. The entire sequence comprises 8 oligonucleotide blocks (1,2,3,4,5,6,17,18) which were synthesized with the phosphamide method. The signal peptide - HBSAg DNA construct, which is claimed, is used to transform a eukaryotic cell, pref. a yeast cell, esp. Saccharomyces cerevisiae AH22R-/pgUD LP39-Rct, L11P39-Rct or LP31-Rct. The advantage is that peptides having HBSAg antigenicity are excised outside the cells and easily purified. The stop codon of the DNA coding for peptide having HBSAg antigenicity may be followed by a terminator (eg PKG terminator) to increase the yield, and a spacer DNA may be inserted between the promoter and the signal peptide DNA. (Updated on 09-JAN-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)

XX Sequence 88 BP; 19 A; 10 C; 23 G; 36 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	826	Length:	88
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	1	Gaps:	0

US-10-030-194A-6 (1-6) x AAN80971 (1-88)

QY 1 GlyTyr\*\*\*ValGlu 5

DB 72 GGTATTTCGTTGAA 86

## RESULT 16

ACH8299/C

ID ACH8299 standard; DNA; 88 BP.

XX





Alignment Scores:  
 Pred. No.: 897 Length: 95  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x AAT30633 (1-95)

QY 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 DB 20 GGATATACAGTGGAA 34

# RESULT 18

AAC67969

ID AAC67969 standard; DNA; 100 BP.

XX

AC AAC67969;

XX 19-FEB-2001 (first entry)

XX Human alpha1-2fucosyltransferase H partial nucleotide sequence #3.

XX Human; alpha1-2fucosyltransferase; cytotatic; Sec1; neuroprotective;

KW nontropic; gene therapy; Fucalpal-2Galbeta1-3GalNAC; immunotherapy;

KW immunosuppression; cancer; neurological disease;

KW small cell lung carcinoma; ds.

XX Homo sapiens.

OS

XX WO200064464-A1.

PN

XX 02-NOV-2000.

PD

XX 23-APR-1999; 99WO-US007384.

PF

XX 23-APR-1999; 99WO-US007384.

PR

XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.

PA

XX Holmes EH, Sherwood AL;

PI

XX WPI; 2000-687262/67.

DR

XX New rat ganglioside GM1-specific alpha1-2Fucosyltransferase, useful for

PT preparation of fucosyl GM1 which is useful as a nutritional composition

PT or immunotherapeutic for cancer and neurological diseases.

PS Example; Fig 1; 91pp; English.

XX

CC The present sequence is given in a specification relating to a rat

CC ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein

CC or its cellular fraction is useful for synthesis of a molecule comprising

CC Fucalpal-2Galbeta1-3GalNAC, a glycolipid, glycoprotein, glycolipoprotein

CC or a free oligosaccharide comprising Fucalpal-2Galbeta1-3GalNAC. The

CC method involves contacting alpha1-2fucosyltransferase with GDP-fucose and

CC a molecule or glycolipid, glycoprotein, glycolipoprotein or

CC oligosaccharide having a terminal Galbeta1-3GalNAC group. It is also

CC useful for synthesis of fucosyl-GM1 by contacting the protein with GDP-

CC fucose and ganglioside GM1. The obtained glycoproteins,

CC glycolipoproteins, glycolipids and oligosaccharides are useful as

CC nutritional compositions and fucosyl-GM1 is useful for inducing an

CC immunotherapeutic or immunosuppressive action against cancer,

CC neurological disease or small cell lung carcinoma

XX

SQ Sequence 100 BP; 23 A; 32 C; 23 G; 22 T; 0 U; 0 Other;

XX

Alignment Scores:

Pred. No.: 948 Length: 100

Score: 22.00 Matches: 4

Percent Similarity: 80.00% Conservative: 0

Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0

DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x AAC67972 (1-100)

XX

Alignment Scores:

Pred. No.: 948 Length: 100

Score: 22.00 Matches: 4

Percent Similarity: 80.00% Conservative: 0

Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x AAC67969 (1-100)

QY 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 DB 11 GGCTACACCGTGGAA 25

# RESULT 19

AAC67972

ID AAC67972 standard; DNA; 100 BP.

XX

AC AAC67972;

XX 19-FEB-2001 (first entry)

XX Human alpha1-2fucosyltransferase Sec2 partial nucleotide sequence #3.

XX Human; alpha1-2fucosyltransferase; cytotatic; Sec2; neuroprotective;

KW nontropic; gene therapy; Fucalpal-2Galbeta1-3GalNAC; immunotherapy;

KW immunosuppression; cancer; neurological disease;

KW small cell lung carcinoma; ds.

XX Homo sapiens.

OS

XX WO200064464-A1.

PN

XX 02-NOV-2000.

PD

XX 23-APR-1999; 99WO-US007384.

PF

XX 23-APR-1999; 99WO-US007384.

PR

XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.

PA

XX Holmes EH, Sherwood AL;

PI

XX WPI; 2000-687262/67.

DR

XX New rat ganglioside GM1-specific alpha1-2Fucosyltransferase, useful for

PT preparation of fucosyl GM1 which is useful as a nutritional composition

PT or immunotherapeutic for cancer and neurological diseases.

PS Example; Fig 1; 91pp; English.

XX

CC The present sequence is given in a specification relating to a rat

CC ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein

CC or its cellular fraction is useful for synthesis of a molecule comprising

CC Fucalpal-2Galbeta1-3GalNAC, a glycolipid, glycoprotein, glycolipoprotein

CC or a free oligosaccharide comprising Fucalpal-2Galbeta1-3GalNAC. The

CC method involves contacting alpha1-2fucosyltransferase with GDP-fucose and

CC a molecule or glycolipid, glycoprotein, glycolipoprotein or

CC oligosaccharide having a terminal Galbeta1-3GalNAC group. It is also

CC useful for synthesis of fucosyl-GM1 by contacting the protein with GDP-

CC fucose and ganglioside GM1. The obtained glycoproteins,

CC glycolipoproteins, glycolipids and oligosaccharides are useful as

CC nutritional compositions and fucosyl-GM1 is useful for inducing an

CC immunotherapeutic or immunosuppressive action against cancer,

CC neurological disease or small cell lung carcinoma

XX

SQ Sequence 100 BP; 23 A; 32 C; 23 G; 22 T; 0 U; 0 Other;

XX

Alignment Scores:

Pred. No.: 948 Length: 100

Score: 22.00 Matches: 4

Percent Similarity: 80.00% Conservative: 0

Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0

DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x AAC67972 (1-100)

XX

Alignment Scores:

Pred. No.: 948 Length: 100

Score: 22.00 Matches: 4

Percent Similarity: 80.00% Conservative: 0

Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0

DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x AAC67972 (1-100)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 5 GGCATGCACTAGA 19  
RESULT 21  
ACD81435  
ID ACD81435 standard; DNA; 100 BP.  
XX ACD81435;  
AC ACD81435;  
XX 19-SEP-2003 (first entry)  
DT 19-SEP-2003 (first entry)  
XX E. coli K12 MG1655 biochip probe SEQ ID 12711.  
DE E. coli K12 MG1655 biochip probe SEQ ID 12711.  
XX Biochip; gene expression; gut; diagnostic; detection; probe; ss.  
KW Escherichia coli.  
XX Escherichia coli.  
OS Escherichia coli.  
XX EP1260592-A1.  
PN EP1260592-A1.  
XX 27-NOV-2002.  
PD 27-NOV-2002.  
XX 17-MAY-2001; 2001EP-00112179.  
PF 17-MAY-2001; 2001EP-00112179.  
XX 17-MAY-2001; 2001EP-00112179.  
PR 17-MAY-2001; 2001EP-00112179.  
XX (MWGB-) MWG-BIOTECH AG.  
PA (MWGB-) MWG-BIOTECH AG.  
XX Donner H, Drescher B, Huber A, Weber J;  
PI WPI; 2003-241155/24.  
XX WPI; 2003-241155/24.  
DR WPI; 2003-241155/24.  
XX Biochip containing probes complementary with open reading frames in  
XX Escherichia coli K12, useful for detecting gene expression and expression  
PT patterns.  
PT patterns.  
XX Claim 3; Page 1979; 2004pp; German.  
XX This invention describes a novel biochip comprising probe spots, each  
CC containing many identical probes. The probes are nucleotide sequences of  
CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at  
CC least one includes a segment of at least 20 bases identical with, or  
CC complementary to, a segment of an open reading frame (orf) of Escherichia  
CC coli K12. The biochip is used for specific detection of gene expression  
CC in K12 and for determining the gene expression pattern, e.g. for  
CC diagnostic determination of which E. coli strains are present in the gut,  
CC and to determine the effects of e.g. growth media on gene expression. The  
CC biochip provides a comprehensive analysis of many different genes with a single  
CC genome, with simultaneous analysis of many different genes with a single  
CC device, and comparison of gene expression between K12 and its mutants or  
CC other E. coli strains in a single experiment. Apart from qualitative and  
CC quantitative information about gene expression, it also allows  
CC measurements of population densities for the various strains. The use of  
CC synthetic oligonucleotides for preparation of probes allows free  
CC variation in probe length and ensures high purity (and thus selectivity,  
CC reactivity and reproducibility); also synthetic probes are generally  
CC shorter than probes prepared by polymerase chain reaction. ACD68731 to  
CC ACD81540 represent oligonucleotide probes used with the biochip described  
CC in the invention  
XX Sequence 100 BP; 24 A; 21 C; 31 G; 24 T; 0 U; 0 Other;  
SQ Sequence 100 BP; 24 A; 21 C; 31 G; 24 T; 0 U; 0 Other;  
Alignment Scores: 948 Length: 100  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 91.67% Indels: 0  
Query Match: 8 Gaps: 0  
DB: 8  
US-10-030-194A-6 (1-6) x ACD81435 (1-100)  
Qy 1 GlyTyr\*\*\*ValGlu 5

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 11 GGCACACCGTGGA 25  
RESULT 20  
ACD81433  
ID ACD81433 standard; DNA; 100 BP.  
XX ACD81433;  
AC ACD81433;  
XX 19-SEP-2003 (first entry)  
DT 19-SEP-2003 (first entry)  
XX E. coli K12 MG1655 biochip probe SEQ ID 12709.  
DE E. coli K12 MG1655 biochip probe SEQ ID 12709.  
XX Biochip; gene expression; gut; diagnostic; detection; probe; ss.  
KW Escherichia coli.  
XX Escherichia coli.  
OS Escherichia coli.  
XX EP1260592-A1.  
PN EP1260592-A1.  
XX 27-NOV-2002.  
PD 27-NOV-2002.  
XX 17-MAY-2001; 2001EP-00112179.  
PF 17-MAY-2001; 2001EP-00112179.  
XX 17-MAY-2001; 2001EP-00112179.  
PR 17-MAY-2001; 2001EP-00112179.  
XX (MWGB-) MWG-BIOTECH AG.  
PA (MWGB-) MWG-BIOTECH AG.  
XX Donner H, Drescher B, Huber A, Weber J;  
PI WPI; 2003-241155/24.  
XX WPI; 2003-241155/24.  
DR WPI; 2003-241155/24.  
XX Biochip containing probes complementary with open reading frames in  
XX Escherichia coli K12, useful for detecting gene expression and expression  
PT patterns.  
PT patterns.  
XX Claim 3; Page 1979; 2004pp; German.  
XX This invention describes a novel biochip comprising probe spots, each  
CC containing many identical probes. The probes are nucleotide sequences of  
CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at  
CC least one includes a segment of at least 20 bases identical with, or  
CC complementary to, a segment of an open reading frame (orf) of Escherichia  
CC coli K12. The biochip is used for specific detection of gene expression  
CC in K12 and for determining the gene expression pattern, e.g. for  
CC diagnostic determination of which E. coli strains are present in the gut,  
CC and to determine the effects of e.g. growth media on gene expression. The  
CC biochip provides a comprehensive analysis of many different genes with a single  
CC genome, with simultaneous analysis of many different genes with a single  
CC device, and comparison of gene expression between K12 and its mutants or  
CC other E. coli strains in a single experiment. Apart from qualitative and  
CC quantitative information about gene expression, it also allows  
CC measurements of population densities for the various strains. The use of  
CC synthetic oligonucleotides for preparation of probes allows free  
CC variation in probe length and ensures high purity (and thus selectivity,  
CC reactivity and reproducibility); also synthetic probes are generally  
CC shorter than probes prepared by polymerase chain reaction. ACD68731 to  
CC ACD81540 represent oligonucleotide probes used with the biochip described  
CC in the invention  
XX Sequence 100 BP; 27 A; 18 C; 23 G; 32 T; 0 U; 0 Other;  
SQ Sequence 100 BP; 27 A; 18 C; 23 G; 32 T; 0 U; 0 Other;  
Alignment Scores: 948 Length: 100  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 91.67% Indels: 0  
Query Match: 8 Gaps: 0  
DB: 8  
US-10-030-194A-6 (1-6) x ACD81433 (1-100)

```

Db          77 GGCTATGCAGTAGAA 91
RESULT 22
ACD81434
ID   ACD81434 standard; DNA; 100 BP.
XX
XX
AC   ACD81434;
XX
XX  19-SEP-2003 (first entry)
XX
XX  E. coli K12 MG1655 biochip probe SEQ ID 12710.
XX
XX  Biochip; gene expression; gut; diagnostic; detection; probe; ss.
XX
XX  Escherichia coli.
XX
XX  EPI260592-A1.
XX
XX  27-NOV-2002.
XX
XX  17-MAY-2001; 2001EP-00112179.
XX
XX  17-MAY-2001; 2001EP-00112179.
XX
XX  (MWGB-) MWG-BIOTECH AG.
XX
XX  Donner H, Drescher B, Huber A, Weber J;
XX
XX  WPI; 2003-241155/24.
XX
XX  Biochip containing probes complementary with open reading frames in
XX  Escherichia coli K12, useful for detecting gene expression and expression
XX  patterns.
XX
XX  Claim 3; Page 1979; 2004pp; German.
XX
XX  This invention describes a novel biochip comprising probe spots, each
XX  containing many identical probes. The probes are nucleotide sequences of
XX  30-80 bases, are prepared ex situ from synthetic oligonucleotides and at
XX  least one includes a segment of at least 20 bases identical with, or
XX  complementary to, a segment of an open reading frame (orf) of Escherichia
XX  coli K12. The biochip is used for specific detection of gene expression
XX  in K12 and for determining the gene expression pattern, e.g. for
XX  diagnostic determination of which E. coli strains are present in the gut,
XX  and to determine the effects of e.g. growth media on gene expression. The
XX  biochip provides as comprehensive as possible detection of the K12
XX  genome, with simultaneous analysis of many different genes with a single
XX  device, and comparison of gene expression between K12 and its mutants or
XX  other E. coli strains in a single experiment. Apart from qualitative and
XX  quantitative information about gene expression, it also allows
XX  measurements of population densities for the various strains. The use of
XX  synthetic oligonucleotides for preparation of probes allows free
XX  variation in probe length and ensures high purity (and thus selectivity,
XX  reactivity and reproducibility); also synthetic probes are generally
XX  shorter than probes prepared by polymerase chain reaction. ACD81540 to
XX  ACD81540 represent oligonucleotide probes used with the biochip described
XX  in the invention
XX
XX  Sequence 100 BP; 23 A; 24 C; 26 G; 27 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.:      948      Length:      100
Score:          22.00     Matches:      4
Percent Similarity: 80.00%  Conservative: 0
Best Local Similarity: 80.00%  Mismatches: 1
Query Match:      91.67%   Indels:      0
DB:               8       Gaps:        0

US-10-030-194A-6 (1-6) x ACD81434 (1-100)
QY  1 GlyTyr***ValGlu 5
    |||||

```

---

```

Db          54 GGCTATGCAGTAGAA 68
RESULT 23
ABS57893
ID   ABS57893 standard; cDNA; 100 BP.
XX
XX  ABS57893;
XX
XX  07-FEB-2003 (first entry)
XX
XX  Human alpha1-2FucT cDNA fragment #6.
XX
XX  Human; ss: alpha1-2fucosyltransferase; alpha1-2FucT; FucT; nutritional;
XX  immunotherapeutic; immunosuppressive; fucosyl-GM_1;
XX  Fucalbeta1-2Galbeta-3GalNAc; cancer; small cell lung carcinoma;
XX  neurological disease; infant formula; geriatric formula; vaccine; RT-PCR;
XX  reverse transcriptase PCR.
XX
XX  Homo sapiens.
XX
XX  US2002127655-A1.
XX
XX  12-SEP-2002.
XX
XX  31-OCT-2001; 2001US-00999672.
XX
XX  23-APR-1999; 99US-00298886.
XX
XX  (HOLM/) HOLMES E H.
XX  (SHER/) SHERWOOD A L.
XX
XX  Holmes EH, Sherwood AL;
XX
XX  WPI; 2003-066903/06.
XX
XX  Novel rat ganglioside GM1-specific alpha 1-2 fucosyltransferase protein
XX  useful in the preparative production of fucosyl-GM1 which is useful as an
XX  immunotherapeutic for cancer.
XX
XX  Example; Fig 1; 44pp; English.
XX
XX  The invention relates to an isolated rat alpha1-2 fucosyltransferase
XX  (FucT) protein appearing as ABG72377 and ABG72378. Also included are: (1)
XX  a chimeric protein comprising FucT, fused by a covalent bond to a
XX  portion of a second protein which is not FucT; (2) an isolated nucleic
XX  acid chosen from the nucleotide sequence of rat hepatoma H35 cell alpha1-
XX  2FucT RT-PCR product and the sequence of the catalytic domain of rat
XX  hepatoma H35 cell alpha1-2FucT; (3) an isolated nucleic acid encoding
XX  FucT, or reverse its complement or RNA equivalent; (4) a vector
XX  comprising the nucleic acid, or a nucleotide sequence that is the reverse
XX  complement to the FucT nucleic acid, and an origin of replication; (5) a
XX  recombinant cell containing the vector; (6) producing FucT; (7) an
XX  isolated and purified protein produced by the above method; (8) a
XX  cellular fraction with protein activity produced by the above method; (9)
XX  detecting the onset of cancer by detecting the nucleotide sequence or its
XX  fragment or complement; (10) suppressing or inhibiting FucT in a cell by
XX  contacting a cell with an antisense RNA corresponding to one of the two
XX  nucleic acid sequences; (11) a nutritional formula composition comprising
XX  glycolipid, glycoprotein, glyco-lipoprotein, or oligosaccharide
XX  synthesised using FucT, the nucleic acid, the chimeric protein, vector
XX  or cell; and (12) inducing an immunotherapeutic or immunosuppressive
XX  action against a fucosyl-GM 1-producing disease, by administering fucosyl
XX  -GM 1 to a human patient with the disease. FucT, the nucleic acid, the
XX  chimeric protein, or the cellular fraction are useful for preparative
XX  synthesis of a molecule comprising Fucalbeta1- 2Galbeta-3GalNAc, by
XX  contacting any one of the above molecules with GDP-fucose and a molecule
XX  having a terminal Galbeta1-3GalNAc group and recovering a molecule
XX  comprising Fucalbeta1-2Galbeta1-3GalNAc, and for preparative synthesis of
XX  a glycolipid, glycoprotein, glyco-lipoprotein or free oligosaccharide
XX  comprising Fucalbeta1-2Galbeta-3GalNAc. The proteins or cellular fraction
XX  are also useful for preparative synthesis of fucosyl-GM 1, by contacting
XX  any of the above protein with GDP-fucose and the ganglioside GM_1 and
XX  recovering fucosyl-GM_1. Inducing an immunotherapeutic or

```

CC immunosuppressive action is useful against a fucosyl-GM 1-producing  
 CC disease such as cancer, especially small cell lung carcinoma or a  
 CC neurological disease. FucT is useful as an immunogen for producing  
 CC antibodies. The glycoproteins, glycolipids, glyco-lipoproteins produced  
 CC by FucT possess nutritional value and are useful as food additives, for  
 CC e.g. infant or geriatric formula. The fucosyl-GM 1 produced using FucT  
 CC serves as a vaccine. The present sequence is a fragment of the human  
 CC alpa1-2FucT cDNA used to design RT-PCR (reverse transcriptase PCR)  
 CC primers used to isolate cDNA encoding rat FucT  
 XX  
 SQ Sequence 100 BP; 23 A; 32 C; 23 G; 22 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 948 Length: 100  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x ABS57893 (1-100)  
 QY 1 GlyTyr\*\*\*ValGlu 5  
 Db 11 GGCTACACCGTGGAA 25

RESULT 24  
 ABS57890  
 ID ABS57890 standard; cDNA; 100 BP.  
 XX AC ABS57890;  
 XX DT 07-FEB-2003 (first entry)  
 XX DE Human alpa1-2FucT cDNA fragment #3.  
 XX KW Human; ss; alpa1-2fucosyltransferase; alpa1-2FucT; FucT; nutritional;  
 KW immunotherapeutic; immunosuppressive; fucosyl-GM 1;  
 KW Fucalpa1-2Galbeta-3GalNAc; cancer; small cell lung carcinoma;  
 KW neurological disease; infant formula; geriatric formula; vaccine; RT-PCR;  
 KW reverse transcriptase PCR.  
 XX KW Homo sapiens.  
 OS  
 XX US2002127655-A1.  
 XX PD 12-SEP-2002.  
 XX PF 31-OCT-2001; 2001US-00999672.  
 XX PR 23-APR-1999; 99US-00298886.  
 XX PA (HOLM/) HOLMES E H.  
 XX PA (SHER/) SHERWOOD A L.  
 XX PI Holmes EH, Sherwood AL;  
 XX WI 2003-066901/06.  
 XX Novel rat ganglioside GM1-specific alpha 1-2 fucosyltransferase protein  
 PT useful in the preparative production of fucosyl-GM1 which is useful as an  
 PT immunotherapeutic for cancer.  
 XX Example; Fig 1; 44pp; English.  
 PS  
 CC The invention relates to an isolated rat alpa1-2 fucosyltransferase  
 CC (FucT) protein appearing as ABG72377 and ABG72378. Also included are: (1)  
 CC a chimeric protein comprising FucT, fused by a covalent bond to a  
 CC portion of a second protein which is not FucT; (2) an isolated nucleic  
 CC acid chosen from the nucleotide sequence of rat hepatoma H35 cell alpa1-  
 CC 2FucT RT-PCR product and the sequence of the catalytic domain of rat  
 CC hepatoma H35 cell alpa1-2FucT; (3) an isolated nucleic acid encoding  
 CC FucT, or reverse its complement or RNA equivalent; (4) a vector

CC comprising the nucleic acid, or a nucleotide sequence that is the reverse  
 CC complement to the FucT nucleic acid, and an origin of replication; (5) a  
 CC recombinant cell containing the vector; (6) producing FucT; (7) an  
 CC isolated and purified protein produced by the above method; (8) a  
 CC cellular fraction with protein activity produced by the above method; (9)  
 CC detecting the onset of cancer by detecting the nucleotide sequence or its  
 CC fragment or complement; (10) suppressing or inhibiting FucT in a cell by  
 CC contacting a cell with an antisense RNA corresponding to one of the two  
 CC nucleic acid sequences; (11) a nutritional formula composition comprising  
 CC glycolipid, glycoprotein, glyco-lipoprotein, or oligosaccharide  
 CC synthesised using FucT, the nucleic acid, the chimeric protein, vector  
 CC or cell; and (12) inducing an immunotherapeutic or immunosuppressive  
 CC action against a fucosyl-GM 1-producing disease, by administering fucosyl  
 CC -GM 1 to a human patient with the disease. FucT, the nucleic acid, the  
 CC chimeric protein, or the cellular fraction are useful for preparative  
 CC synthesis of a molecule comprising Fucalpa1-2Galbeta-3GalNAc, by  
 CC contacting any one of the above molecules with GDP-fucose and a molecule  
 CC having a terminal Galbeta1-3GalNAc group and recovering a molecule  
 CC comprising Fucalpa1-2Galbeta1-3GalNAc, and for preparative synthesis of  
 CC a glycolipid, glycoprotein, glyco-lipoprotein or free oligosaccharide  
 CC comprising Fucalpa1-2Galbeta-3GalNAc. The proteins or cellular fraction  
 CC are also useful for preparative synthesis of fucosyl-GM 1, by contacting  
 CC any of the above protein with GDP-fucose and the ganglioside GM 1 and  
 CC recovering fucosyl-GM 1. Inducing an immunotherapeutic or  
 CC immunosuppressive action is useful against a fucosyl-GM 1-producing  
 CC disease such as cancer, especially small cell lung carcinoma or a  
 CC neurological disease. FucT is useful as an immunogen for producing  
 CC antibodies. The glycoproteins, glycolipids, glyco-lipoproteins produced  
 CC by FucT possess nutritional value and are useful as food additives, for  
 CC e.g. infant or geriatric formula. The fucosyl-GM 1 produced using FucT  
 CC serves as a vaccine. The present sequence is a fragment of the human  
 CC alpa1-2FucT cDNA used to design RT-PCR (reverse transcriptase PCR)  
 CC primers used to isolate cDNA encoding rat FucT  
 XX  
 SQ Sequence 100 BP; 23 A; 32 C; 23 G; 22 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 948 Length: 100  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x ABS57890 (1-100)

QY 1 GlyTyr\*\*\*ValGlu 5

Db 11 GGCTACACCGTGGAA 25

RESULT 25

AAF75449  
 ID AAF75449 standard; DNA; 108 BP.

XX AC AAF75449;

XX DT 14-MAY-2001 (first entry)

XX DE Codon-optimised HPV16 E2 fragment 13856-307-2E.

XX KW Human papillomavirus; HPV; HPV16; HPV6a; HPV18; L1; E2; E7; E1;  
 XX antiviral; immunostimulant; vaccine; immunogen; infection; ss.

XX OS Human papillomavirus.

XX OS Synthetic.

XX PN WO200114416-A2.

XX PD 01-MAR-2001.

XX PF 21-AUG-2000; 2000WO-US022932.

XX PR 25-AUG-1999; 99US-0150728P.

```

PR 07-JUN-2000; 2000US-0210143P.
PA (MERI ) MERCK & CO INC.
XX
XX Nepper MP, McClements WL, Jansen KU, Schultz LD, Chen L, Wang X;
XX WPI; 2001-218428/22.
XX
XX Novel synthetic polynucleotide encoding human papillomavirus (HPV)
PT protein or mutated HPV protein useful as anti-HPV vaccines, comprises
PT optimized-codons for expression of the viral proteins in human host
PT cells.
XX
XX Example 4; Fig 19; 119pp; English.
XX
XX The present sequence is an oligomer which was used in the assembly of one
CC of a number of synthetic polynucleotides that encode a human
CC papillomavirus (HPV) protein, or a mutated form of a HPV protein. The
CC mutated HPV proteins have reduced protein function as compared to wild
CC type proteins but maintain immunogenicity. The proteins comprise codons
CC for optimised expression in humans. The polynucleotides are useful as a
CC vaccine which provides effective immunoprophylaxis against papillomavirus
CC infection through stimulation of neutralising antibody and cell-mediated
CC immunity
XX
SQ Sequence 108 BP; 29 A; 33 C; 28 G; 18 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.03e+03 Length: 108
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x AAF75449 (1-108)

Qy 1 GlyTyr***ValGlu 5
Db 21 GGCTACACCGTGGAG 35

RESULT 26
ADK93589
ID ADK93589 standard; DNA; 121 BP.
XX
AC ADK93589;
XX
XX 06-MAY-2004 (first entry)
XX
XX Polynucleotide used to detect SNPs of the invention #2618.
XX human; single nucleotide polymorphism; SNP; ds.
XX
XX Homo sapiens.
XX
XX JP2003259875-A.
XX
XX 16-SEP-2003.
XX
XX 08-MAR-2002; 2002JP-00064373.
XX
XX 08-MAR-2002; 2002JP-00064373.
XX
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
XX WPI; 2004-093977/10.
XX
XX Novel polynucleotide useful for PCR amplification along with two DNA
PT fragment from another set of sequences, or for detecting single
PT nucleotide polymorphism in human gene.
XX
XX Claim 1; SEQ ID NO 2618; 2627pp; Japanese.
XX

```

---

```

CC The present invention relates to a polynucleotide isolated from a human
CC gene and is useful for detecting a single nucleotide polymorphism in a
CC human gene or for diagnosing of disease. The invention enables the
CC detection of a single nucleotide polymorphism in a human gene. The
CC present sequence represents a polynucleotide used to detect SNPs of the
CC invention.
XX
XX Sequence 121 BP; 31 A; 14 C; 24 G; 51 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 1.17e+03 Length: 121
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 12 Gaps: 0

US-10-030-194A-6 (1-6) x ADK93589 (1-121)

Qy 1 GlyTyr***ValGlu 5
Db 43 GGCTATTCTGTGGAG 57

RESULT 27
ACH84646/C
ID ACH84646 standard; DNA; 124 BP.
XX
AC ACH84646;
XX
XX 29-JUL-2004 (first entry)
XX
XX Human genome derived single exon probe #17841.
XX
XX Human; probe; ss; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.
XX
XX Homo sapiens.
XX
XX US2003194704-A1.
XX
XX 16-OCT-2003.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX (PENN/) PENN S G.
XX (RANK/) RANK D R.
XX (HANZ/) HANZEL D K.
XX
XX Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX Claim 1; SEQ ID NO 17841; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of

```

measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, a customer desiring to measure gene expression, a method of providing human gene expression data by a database having a plurality of records storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising gross alterations in the genomic locus that includes the synthesis of nucleic acids, smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030194704

Sequence 124 BP; 24 A; 35 C; 36 G; 29 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.2e+03 Length: 124  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 12 Gaps: 0

US-10-030-194A-6 (1-6) x ACH84646 (1-124)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 43 GGCTACAGCGTGGAA 29

RESULT 28  
ABN80737  
ID ABN80737 standard; DNA; 126 BP.  
XX AC ABN80737;  
XX DT 15-JUL-2002 (first entry)  
XX DE Unique bacterial gene dhfr2 from clone pASDMN1 used as a probe.  
XX KW Primer; probe; affinity purification; cloning; biosynthetic pathway;  
KW hybridization screening; restriction endonuclease; aac6'; acvs; dhfr2;  
KW ipns; NIM; PKS; strB1; strD; strE; strA; stcC; taxol; ss.  
XX OS Unidentified.  
XX PN WO200112861-A1.  
XX PD 22-FEB-2001.  
XX PF 18-AUG-2000; 2000WO-US022743.  
XX PR 19-AUG-1999; 99US-0149788P.  
PR 19-SEP-1999; 99US-0149822P.  
XX XX (OMNI-) OMNISCIENCE PHARM.  
XX XX Fomenkov A, Huang Y, Chaparian MG, Zheng S;  
XX PI WPT; 2001-211237/21.  
XX DR New primers and probes useful for targeted cloning and enrichment of  
XX PT genes and gene clusters for affinity purification of genes or in cloning  
PT associated biosynthetic pathway genes.

associated biosynthetic pathway genes.  
XX Claim 11; Page 42-43; 74pp; English.  
XX The invention relates to a set of novel primers and probes. The genes  
CC cloned are used in affinity purification of genes, and for cloning  
CC associated biosynthetic pathway genes. The gene probes/primers may be  
CC used in the discovery and characterization of bioactive compound coding  
CC sequences and gene clusters, as well as in the discovery of either single  
CC genes or entire clusters of adjacent genes involved in the total  
CC synthesis of compounds of interest, e.g. secondary metabolite  
CC biosynthetic pathways the products of which comprise very useful  
CC libraries for antibiotic and other therapeutic compound screening. The  
CC cloned genes are further useful for hybridization screening. The sequence  
CC represents a unique cloned gene sequence used as a probe in the invention  
XX Sequence 126 BP; 24 A; 38 C; 37 G; 27 T; 0 U; 0 Other;  
SQ Alignment Scores:  
Pred. No.: 1.22e+03 Length: 126  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x ABN80737 (1-126)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 78 GGCTATGCCGTCGAG 92

RESULT 29  
ABN80739/c  
ID ABN80739 standard; DNA; 127 BP.  
XX AC ABN80739;  
XX DT 15-JUL-2002 (first entry)  
XX DE Unique bacterial gene dhfr2 from clone pASDMN5 used as a probe.  
XX KW Primer; probe; affinity purification; cloning; biosynthetic pathway;  
KW hybridization screening; restriction endonuclease; aac6'; acvs; dhfr2;  
KW ipns; NIM; PKS; strB1; strD; strE; strA; stcC; taxol; ss.  
XX OS Unidentified.  
XX PN WO200112861-A1.  
XX PD 22-FEB-2001.  
XX PF 18-AUG-2000; 2000WO-US022743.  
XX PR 19-AUG-1999; 99US-0149788P.  
PR 19-SEP-1999; 99US-0149822P.  
XX XX (OMNI-) OMNISCIENCE PHARM.  
XX XX Fomenkov A, Huang Y, Chaparian MG, Zheng S;  
XX PI WPT; 2001-211237/21.  
XX DR New primers and probes useful for targeted cloning and enrichment of  
XX PT genes and gene clusters for affinity purification of genes or in cloning  
PT associated biosynthetic pathway genes.  
XX Claim 11; Page 42-43; 74pp; English.  
XX The invention relates to a set of novel primers and probes. The genes  
CC cloned are used in affinity purification of genes, and for cloning  
CC associated biosynthetic pathway genes. The gene probes/primers may be  
CC used in the discovery and characterization of bioactive compound coding  
CC used in the discovery and characterization of bioactive compound coding

CC sequences and gene clusters, as well as in the discovery of either single  
CC genes or entire clusters of adjacent genes involved in the total  
CC synthesis of compounds of interest, e.g. secondary metabolite  
CC biosynthetic pathways the products of which comprise very useful  
CC libraries for antibiotic and other therapeutic compound screening. The  
CC cloned genes are further useful for hybridization screening. The sequence  
CC represents a unique cloned gene sequence used as a probe in the invention  
XX  
SQ Sequence 127 BP; 26 A; 38 C; 36 G; 27 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.23e+03 Length: 127  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x ABN80739 (1-127)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 49 GGCTATGCCGTCGAG 35

RESULT 30

ABN80738/c  
ID ABN80738 standard; DNA; 127 BP.

AC ABN80738;

DT 15-JUL-2002 (first entry)

DE Unique bacterial gene dhfr2 from clone PASDMN2 used as a probe.

KW Primer; probe; affinity purification; cloning; biosynthetic pathway;  
KW hybridization screening; restriction endonuclease; aac6'; acvs; dhfr2;  
KW ipns; NIM; PKS; strB1; strD; strE; stcA; stcC; taxol; ss.

XX Unidentified.

XX WO200112861-A1.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-US022743.

XX 19-AUG-1999; 99US-0149789P.

XX 19-SEP-1999; 99US-0149822P.

XX (OMNI-) OMNISCIENCE PHARM.

XX Fomenkov A, Huang Y, Chaparian MG, Zheng S;

XX WPI; 2001-211237/21.

XX New primers and probes useful for targeted cloning and enrichment of  
PT genes and gene clusters for affinity purification of genes or in cloning  
PT associated biosynthetic pathway genes.

XX Claim 11; Page 42-43; 74pp; English.

XX The invention relates to a set of novel primers and probes. The genes  
CC cloned are used in affinity purification of genes, and for cloning  
CC associated biosynthetic pathway genes. The gene probes/primers may be  
CC used in the discovery and characterization of bioactive compound coding  
CC sequences and gene clusters, as well as in the discovery of either single  
CC genes or entire clusters of adjacent genes involved in the total  
CC synthesis of compounds of interest, e.g. secondary metabolite  
CC biosynthetic pathways the products of which comprise very useful  
CC libraries for antibiotic and other therapeutic compound screening. The  
CC cloned genes are further useful for hybridization screening. The sequence  
CC represents a unique cloned gene sequence used as a probe in the invention  
XX

SQ Sequence 127 BP; 28 A; 37 C; 38 G; 24 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.23e+03 Length: 127  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x ABN80738 (1-127)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 49 GGCTATGCCGTCGAG 35

RESULT 31

ABA73680/c

ID ABA73680 standard; DNA; 136 BP.

XX AC ABA73680;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #21985.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human fetal liver.  
XX Claim 4; SEQ ID NO 21985; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring  
CC human gene expression in a sample derived from human foetal liver. The  
CC single exon nucleic acid probes may be used for predicting, measuring and  
CC displaying gene expression in samples derived from human fetal liver. The  
CC present sequence is a single exon nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 136 BP; 31 A; 29 C; 28 G; 48 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.32e+03 Length: 136  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x ABA73680 (1-136)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 127 GGTATTCCGTGGAA 113

RESULT 32  
AAI54119/C  
ID AAI54119 standard; DNA; 136 BP.

AC AAI54119;

XX 17-OCT-2001 (first entry)

XX Probe #22805 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;

KW genetic disorder; ss.

XX Homo sapiens.

OS

XX WO200157272-A2.

PN

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta.

XX Claim 25; SEQ ID NO 22805; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP).

CC The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders

XX Sequence 136 BP; 31 A; 29 C; 28 G; 48 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.32e+03 Length: 136  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x AAI54119 (1-136)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 127 GGTATTCCGTGGAA 113

RESULT 33  
AAK48294/C  
ID AAK48294 standard; DNA; 136 BP.

XX

AC AAK48294;

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 22851.

DE Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

XX WO200157276-A2.

PN

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000668.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human bone marrow.

XX Example 4; SEQ ID NO 22851; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukaemia and myeloma. The present sequence is one of the probes of the invention

XX Sequence 136 BP; 31 A; 29 C; 28 G; 48 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.32e+03 Length: 136  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x AAK48294 (1-136)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 127 GGTATTCCGTGGAA 113

RESULT 34  
AAK22127/C

ID AAK22127 standard; DNA; 136 BP.

XX

XX AAK22127;

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe SEQ ID NO: 22118.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;

KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;

XX ss.



OS Homo sapiens.  
 PN WO200157275-A2.  
 XX  
 XX  
 PD  
 XX  
 XX  
 PF 30-JAN-2001; 2001WO-US000667.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-483446/52.  
 DR  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains.  
 XX  
 XX Example 4; SEQ ID NO 22118; 650pp + Sequence Listing; English.  
 PS  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is one of the probes of the  
 CC invention  
 XX  
 XX Sequence 136 BP; 31 A; 29 C; 28 G; 48 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.32e+03 Length: 136  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-030-194A-6 (1-6) x AAK22127 (1-136)  
 Qy 1 GlyTyr\*\*\*ValGlu 5  
 Db 127 GGTATTCCGTGGAA 113  
 RESULT 35  
 ABS47995/C  
 ID ABS47995 standard; DNA; 136 BP.  
 AC  
 XX ABS47995;  
 XX  
 XX 25-FEB-2003 (first entry)  
 DT  
 XX Human liver single exon probe, SEQ ID NO 22985.  
 DE  
 XX Human; single exon nucleic acid probe; liver; cirrhosis;  
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;  
 KW coronary heart disease; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200157273-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US000664.  
 PF  
 XX

PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488898/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human adult liver.  
 PT  
 XX Claim 4; SEQ ID NO 22985; 658pp; English.  
 PS  
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult liver.  
 CC (I) may be used for predicting, measuring and displaying gene expression  
 CC in samples derived from human adult liver. The genes identified may be  
 CC involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
 CC associated with coronary heart disease. ABS25011-ABS51005 represent human  
 CC liver single exon nucleic acid probes of the invention. Note: The  
 CC sequence information for this patent does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 136 BP; 31 A; 29 C; 28 G; 48 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1.32e+03 Length: 136  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-030-194A-6 (1-6) x ABS47995 (1-136)  
 Qy 1 GlyTyr\*\*\*ValGlu 5  
 Db 127 GGTATTCCGTGGAA 113  
 RESULT 36  
 AAV79014  
 ID AAV79014 standard; DNA; 166 BP.  
 XX  
 XX AAV79014;  
 AC  
 XX  
 XX 16-MAR-1999 (first entry)  
 DT  
 XX Staphylococcus aureus contig SEQ ID #4703.  
 DE  
 XX Computer readable medium; vaccine; S.aureus infection; immunodetection;  
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
 KW skin infection; surgical wound infection; scalded skin syndrome;  
 KW toxic shock syndrome; ds.  
 XX  
 XX Staphylococcus aureus.  
 OS  
 XX EP786519-A2.  
 PN  
 XX 30-JUL-1997.  
 PD  
 XX 07-JAN-1997; 97EP-00100117.  
 PF  
 XX

Thu Nov 4 17:32:32 2004

us-10-030-194a-6.p2n.rng

PR 05-JAN-1996; 96US-0009861P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Kunsch CA, Choi GH, Barash SC, Dillion PU, Fannon MR, Rosen CA;  
 FI WPI; 1997-374922/35.  
 XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -  
 PT stored on computer readable medium and used in the production of anti-  
 PT S.aureus vaccines.  
 XX Claim 1; Page 3059; 3271pp; English.  
 PS This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
 CC of the invention. The DNA sequences are recorded on a computer readable  
 CC medium, preferably selected from a floppy or hard disk, random access  
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
 CC the S.aureus DNA sequences allows putative functions to be assigned so  
 CC that protein-encoding or regulatory regions of commercial, therapeutic or  
 CC industrial importance can be obtained. Specifically, sequences which are  
 CC likely to encode antigens have been identified and these polypeptides can  
 CC be used in a vaccine composition against S.aureus infection. The  
 CC polypeptides can also be used in a kit for the immunodetection of  
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,  
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
 CC for recombinant production of the polypeptides. The new DNA sequences  
 CC (and their fragments) are useful as primers or probes for isolating  
 CC homologues of any of the S.aureus DNA sequences contained on the computer  
 CC readable medium  
 XX Sequence 166 BP; 49 A; 27 C; 41 G; 48 T; 0 U; 1 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.64e+03 Length: 166  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 2 Gaps: 0  
 US-10-030-194A-6 (1-6) x AAV79014 (1-166)  
 QY 1 GlyTyr\*\*\*ValGlu 5  
 DB 13 GGTATTCTGTGAA 27  
 RESULT 37  
 ABN16095  
 ID ABN16095 standard; cDNA; 169 BP.  
 XX AC ABN16095;  
 XX 24-JUN-2002 (first entry)  
 DT Human ORFX polynucleotide sequence SEQ ID NO:667.  
 DE  
 XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 KW hypertension; hypothyroidism; cholesterol ester storage disease;  
 KW immune deficiency; immune disorder; infectious disease;  
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 KW myasthenia gravis; Gene; ss.  
 XX Homo sapiens.  
 OS WO200192523-A2.  
 PN 06-DEC-2001.  
 XX

XX 29-MAY-2001; 2001WO-US010836.  
 PF 30-MAY-2000; 2000US-0206132P.  
 PR 29-AUG-2000; 2000US-0228716P.  
 XX (CURA-) CURAGEN CORP.  
 PA Shinkets RA, Leach MD;  
 XX WPI; 2002-106308/14.  
 PI P-PSDB; ABP00343.  
 DR Novel human polypeptides and polynucleotides useful for diagnosing,  
 XX preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders.  
 PS Disclosure; SEQ ID NO 667; 1037pp; English.  
 XX The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX  
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
 CC treating or preventing a pathology associated with an ORFX-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
 CC sequences can be used in gene therapy. ORFX sequences can be used in the  
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
 CC osteoarthritis, neurodegenerative diseases, diabetes mellitus, systemic  
 CC transplantation, cardiovascular diseases, disorders related to organ  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, rheumatoid  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 169 BP; 58 A; 31 C; 29 G; 48 T; 0 U; 3 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.67e+03 Length: 169  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-030-194A-6 (1-6) x ABN16095 (1-169)  
 QY 1 GlyTyr\*\*\*ValGlu 5  
 DB 126 GGTATTCTGTGAG 140  
 RESULT 38  
 ABA69543/c  
 ID ABA69543 standard; DNA; 189 BP.  
 XX AC ABA69543;  
 XX 01-FEB-2002 (first entry)  
 DT Human foetal liver single exon nucleic acid probe #17848.  
 DE Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
 XX Homo sapiens.  
 OS

XX PN WO200157277-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000669.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483447/52.  
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing  
XX PT gene expression in human fetal liver.  
XX PS Claim 4; SEQ ID NO 17848; 639pp + Sequence Listing; English.  
XX CC The invention relates to a single exon nucleic acid probe for measuring  
XX CC human gene expression in a sample derived from human foetal liver. The  
XX CC single exon nucleic acid probes may be used for predicting, measuring and  
XX CC displaying gene expression in samples derived from human fetal liver. The  
XX CC present sequence is a single exon nucleic acid probe of the invention.  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 189 BP; 49 A; 48 C; 36 G; 55 T; 0 U; 1 Other;  
Alignment Scores:  
Pred. No.: 1.89e+03 Length: 189  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0  
US-10-030-194A-6 (1-6) x ABA69543 (1-189)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 138 GGCTATGCTGTAGAA 124  
RESULT 39  
ABA36466/C  
ID ABA36466 standard; DNA; 189 BP.  
XX AC ABA36466;  
XX DT 23-JAN-2002 (first entry)  
XX DE Probe #14932 for gene expression analysis in human heart cell sample.  
XX KW Human; gene expression; heart; microarray; vascular system; probe;  
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;  
XX KW congenital heart disease; ss.  
XX OS Homo sapiens.  
XX PN WO200157274-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000666.  
XX

PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488899/53.  
XX DR Single exon nucleic acid probes for analyzing gene expression in human  
XX PT hearts.  
XX PS Claim 4; SEQ ID NO 14932; 530pp; English.  
XX CC The present invention relates to single exon nucleic acid probes for  
XX CC measuring human gene expression in a sample derived from human heart. The  
XX CC present sequence is one such probe. The probes may be used for  
XX CC predicting, measuring and displaying gene expression in samples derived  
XX CC from the human heart via microarrays. By measuring gene expression, the  
XX CC probes are useful for predicting, diagnosing, grading, staging,  
XX CC monitoring and prognosing diseases of the human heart and vascular system  
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
XX CC congenital heart disease. Note: The sequence data for this patent did not  
XX CC form part of the printed specification, but was obtained in electronic  
XX CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 189 BP; 49 A; 48 C; 36 G; 55 T; 0 U; 1 Other;  
Alignment Scores:  
Pred. No.: 1.89e+03 Length: 189  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0  
US-10-030-194A-6 (1-6) x ABA36466 (1-189)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 138 GGCTATGCTGTAGAA 124  
RESULT 40  
AAI09933/C  
ID AAI09933 standard; DNA; 189 BP.  
XX AC AAI09933;  
XX DT 09-OCT-2001 (first entry)  
XX DE Probe #9924 used to measure gene expression in human breast sample.  
XX KW Probe; human; breast disease; breast cancer; development disorder; ss;  
XX KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX OS Homo sapiens.  
XX PN WO200157270-A2.  
XX PD 09-AUG-2001.  
XX PF 29-JAN-2001; 2001WO-US000661.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX

PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-476286/51.  
XX Novel single exon nucleic acid probe used to measuring gene expression in  
XX a human breast.  
XX Claim 25; SEQ ID NO 9924; 322pp; English.  
XX The present invention relates to novel single exon nucleic acid probes.  
XX The present sequence is one such probe. The probes are useful for  
XX measuring human gene expression in a human breast sample, where the probe  
XX hybridises at high stringency to a nucleic acid expressed in the human  
XX breast. The probes are useful for predicting, diagnosing, grading,  
XX staging, monitoring and prognosing diseases of the human breast,  
XX particularly those diseases with polygenic aetiology. The diseases  
XX include: breast cancer, disorders of development, inflammatory diseases  
XX of the breast, fibrocystic changes, proliferative breast disease and non-  
XX carcinoma tumours. Note: The sequence data for this patent did not form  
XX part of the printed specification, but was obtained in electronic format  
XX directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 189 BP; 49 A; 48 C; 36 G; 55 T; 0 U; 1 Other;  
XX  
XX Alignment Scores:  
XX Pred. No.: 1.89e+03 Length: 189  
XX Score: 22.00 Matches: 4  
XX Percent Similarity: 80.00% Conservative: 0  
XX Best Local Similarity: 80.00% Mismatches: 1  
XX Query Match: 91.67% Indels: 0  
XX DB: 5 Gaps: 0  
XX  
XX US-10-030-194A-6 (1-6) x AA109933 (1-189)  
XX  
XX QY 1 GlyTyr\*\*\*ValGlu 5  
XX 138 GCCTATGCTGTAGAA 124  
XX  
XX DB  
XX  
XX RESULT 41  
XX ABS17782/c  
XX ID ABS17782 standard; DNA; 189 BP.  
XX  
XX AC ABS17782;  
XX  
XX CC 19-AUG-2002 (first entry)  
XX  
XX DE Human genome-derived single exon probe ORF from lung SEQ ID No 17773.  
XX  
XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
XX chronic obstructive pulmonary disease; interstitial lung disease;  
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;  
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
XX primary ciliary dyskinesia; pulmonary hypertension;  
XX hyaline membrane disease; open reading frame; ORF.  
XX  
XX OS Homo sapiens.  
XX  
XX FN WO200186003-A2.  
XX  
XX PD 15-NOV-2001.  
XX  
XX PF 30-JAN-2001; 2001WO-US000665.  
XX  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX

PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2002-114183/15.  
XX Spatially-addressable set of single exon nucleic acid probes, used to  
XX measure gene expression in human lung samples.  
XX Claim 4; SEQ ID NO 17773; 634pp; English.  
XX The invention relates to a spatially-addressable set of single exon  
XX nucleic acid probes for measuring gene expression in a sample derived  
XX from human lung comprising single exon nucleic acid probes having one of  
XX 12614 nucleic acid sequences mentioned in the specification, or their  
XX complements or the 12387 open reading frames derived from the 12614  
XX probes. Also included are a microarray comprising the novel set of probes  
XX; the novel set of probes which hybridise at high stringency to a nucleic  
XX acid expressed in the human lung; measuring gene expression in a sample  
XX derived from human lung, comprising (a) contacting the array with a  
XX collection of detectably labeled nucleic acids derived from human lung  
XX mRNA, and (b) measuring the label detectably bound to each probe of the  
XX array; identifying exons in a eukaryotic genome, comprising (a)  
XX algorithmically predicting at least one exon from genomic sequences of  
XX the eukaryote; and (b) detecting specific hybridisation of detectably  
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
XX having a fragment identical to the predicted exon, the probe is included  
XX in the above mentioned microarray; assigning exons to a single gene,  
XX comprising (a) identifying exons from genomic sequence by the method  
XX above and (b) measuring the expression of each of the exons in several  
XX tissues and/or cell types using hybridisation to a single exon  
XX microarrays having a probe with the exon where a common pattern of  
XX expression of the exons in the tissues and/or cell types indicates that  
XX the exons should be assigned to a single gene; a peptide comprising one  
XX of 12011 sequences, mentioned in the specification, or encoded by the  
XX probes/open reading frames (ORF). The probes are used for gene expression  
XX analysis, and for identifying exons in a gene, particularly using human  
XX lung derived mRNA and for the study of lung diseases such as asthma, lung  
XX cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
XX disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
XX tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
XX Pudlak syndrome, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
XX histiocytosis, fibrocystic pulmonary dysplasia, primary ciliary  
XX Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
XX dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
XX present sequence is a single exon probe open reading frame of the  
XX invention. Note: The sequence data for this patent did not form part of  
XX the printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 189 BP; 49 A; 48 C; 36 G; 55 T; 0 U; 1 Other;  
XX  
XX Alignment Scores:  
XX Pred. No.: 1.89e+03 Length: 189  
XX Score: 22.00 Matches: 4  
XX Percent Similarity: 80.00% Conservative: 0  
XX Best Local Similarity: 80.00% Mismatches: 1  
XX Query Match: 91.67% Indels: 0  
XX DB: 6 Gaps: 0  
XX  
XX US-10-030-194A-6 (1-6) x ABS17782 (1-189)  
XX  
XX QY 1 GlyTyr\*\*\*ValGlu 5  
XX 138 GCCTATGCTGTAGAA 124  
XX  
XX DB  
XX  
XX RESULT 42

AAD37663/c  
ID AAD37663 standard; DNA; 193 BP.  
XX  
AC AAD37663;  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE Maize ATP synthase beta subunit MTP DNA.  
XX  
KW Amino acid oxidase; insect infestation; gene therapy; boll weevil; BWV;  
KW corn rootworm; CRW; insecticide; wireworm; WW; Colorado potato beetle;  
KW CPB; maize; ATP synthase; enzyme; mitochondrial targeting peptide; MTP;  
KW ds.  
XX  
OS Zea mays.  
XX  
PN US6372211-B1.  
XX  
PD 16-APR-2002.  
XX  
PF 21-APR-1998; 98US-00063733.  
XX  
PR 21-APR-1997; 97US-0044504P.  
XX  
PA (MONS ) MONSANTO TECHNOLOGY LLC.  
XX  
PI Isaac BG, Greenplate JT, Purcell JP, Romano CP;  
XX  
DR WPI; 2002-424731/45.  
XX  
PT Composition for controlling coleopteran insect infestation of plants,  
PT such as Diabrotica, Melanotus, Leptinotarsa, or Anthonomus infestation,  
PT comprises a lysine oxidase enzyme and a tadanalactam synthase enzyme.  
XX  
PS Example 11; Col 81-82; 77pp; English.  
XX  
CC The invention relates to composition and methods for controlling  
CC coleopteran insect infestation of plants, by co-expressing an amino acid  
CC oxidase and a second enzyme that provides insecticidal activity when  
CC present in a mixture with the amino acid oxidase. Nucleic acid sequences  
CC encoding these enzymes are used in gene therapy. The composition is used  
CC to control insect infestation of plants. It is used to control  
CC coleopteran species Diabrotica, Anthonomus, Melanotus, or Leptinotarsa.  
CC It is especially used to control boll weevil (BWV), corn rootworm (CRW),  
CC wireworm (WW) or Colorado potato beetle (CPB). The composition can also  
CC be used to transform plants that can express the enzymes of the  
CC composition. The present sequence is maize ATP synthase beta subunit  
CC mitochondrial targeting peptide (MTP) DNA. This sequence is used in the  
CC exemplification of the invention.  
XX  
SQ Sequence 193 BP; 19 A; 92 C; 46 G; 36 T; 0 U; 0 Other;  
  
Alignment Scores: 1.93e+03 Length: 193  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 6  
  
US-10-030-194A-6 (1-6) x AAD37663 (1-193)  
  
OY 1 GlyTyr\*\*\*ValGlu 5  
DB 147 GGCTACACGGTTGAG 133  
  
RESULT 43  
ID ACD29603/c  
AC ACD29603 standard; DNA; 193 BP.  
XX  
AC ACD29603;  
XX  
DT 28-AUG-2003 (first entry)  
XX

XX Maize ATP synthase beta subunit mitochondrial targeting peptide DNA.  
DE  
XX  
KW Maize; ATP synthase; MTP; mitochondrial targeting peptide; fungus; ds;  
KW plant antifungal; insect control; boll weevil; corn rootworm; wireworm;  
KW Colorado potato beetle.  
XX  
OS Zea mays.  
XX  
PN US2003026795-A1.  
XX  
PD 06-FEB-2003.  
XX  
PF 26-OCT-2001; 2001US-00005530.  
XX  
PR 21-APR-1997; 97US-0044504P.  
PR 21-APR-1998; 98US-00063733.  
XX  
PA (MONS ) MONSANTO TECHNOLOGY LLC.  
XX  
PI Isaac BG, Greenplate JT, Purcell JP, Romano CP;  
XX  
DR WPI; 2003-492026/46.  
XX  
PT Composition for controlling insect infestation of plants, comprises a  
PT mixture of enzymes comprising amino acid oxidase, and provides  
PT insecticidal activity.  
XX  
PS Example 11; Page 43-44; 86pp; English.  
XX  
CC The invention relates to a composition for controlling insect infestation  
CC of plants which comprises a mixture of a first enzyme comprising an amino  
CC acid oxidase and a second enzyme that provides insecticidal activity when  
CC present in the mixture with the first enzyme, where the mixture is  
CC ingested by an insect. The composition is useful for controlling insect  
CC infestation of plants by boll weevil, corn rootworm, wireworm and  
CC Colorado potato beetle. The present sequence represents maize ATP  
CC synthase beta subunit mitochondrial targeting peptide MTP DNA  
XX  
SQ Sequence 193 BP; 19 A; 92 C; 46 G; 36 T; 0 U; 0 Other;  
  
Alignment Scores: 1.93e+03 Length: 193  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 8  
  
US-10-030-194A-6 (1-6) x ACD29603 (1-193)  
  
OY 1 GlyTyr\*\*\*ValGlu 5  
DB 147 GGCTACACGGTTGAG 133  
  
RESULT 44  
ABZ18648  
ID ABZ18648 standard; cDNA; 202 BP.  
XX  
AC ABZ18648;  
XX  
DT 23-JAN-2003 (first entry)  
XX  
DE Group III cDNA cancer related clone SEQ ID NO:1074.  
XX  
KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;  
KW immune response; virology; immunology; microbiology; molecular biology;  
KW recombinant DNA technology; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2000278516-A2.  
XX

PD 10-OCT-2002.  
 XX 28-MAR-2002; 2002WO-US010421.  
 PF 30-MAR-2001; 2001US-0280255P.  
 PR 28-AUG-2001; 2001US-0315533P.  
 PR 09-JAN-2002; 2002US-0347313P.  
 XX (CORI-) CORIXA CORP.  
 PA Wang T, Wang S, Bangur CS, Gaiger A;  
 XX WPI; 2003-058387/05.  
 XX New immunogenic polynucleotides or polypeptides useful for diagnosing, in  
 PT preventing and treating cancer expressing CT or CP mRNA antigens, and in  
 PT virology, immunology, microbiology, molecular biology and recombinant DNA  
 PT techniques.  
 XX Claim 1; SEQ ID NO 1074; 207pp; English.  
 PS ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and  
 CC ABP54446 to ABP54472 represent protein (II) sequences, from the present  
 CC invention; (I) and (II) have cytostatic activity and can be used in gene  
 CC therapy and vaccines. (I), (II), antibodies and compositions from the  
 CC present invention are useful for diagnosing, preventing and treating  
 CC cancer, which expresses CT or CP mRNA antigens. They are useful for  
 CC stimulating immune response. They can also be useful in virology,  
 CC immunology, microbiology, molecular biology and recombinant DNA  
 CC techniques. N.B. The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 202 BP; 44 A; 55 C; 61 G; 42 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores: Length: 202  
 Pred. No.: 2.03e+03 Matches: 4  
 Score: 22.00 Conservative: 0  
 Percent Similarity: 80.00% Mismatches: 1  
 Best Local Similarity: 80.00% Indels: 0  
 Query Match: 91.67% Gaps: 0  
 DB: 8  
 US-10-030-194A-6 (1-6) x ABZ18648 (1-202)  
 QY 1 GlyTyr\*\*\*ValGlu 5  
 DB 149 GGCTATGCTGTGGAG 163  
 RESULT 45  
 ACAS5687/C  
 ID ACAS5687 standard; cDNA; 202 BP.  
 XX ACAS5687;  
 AC 06-JUN-2003 (first entry)  
 XX Rabbit signalling pathway polynucleotide probe SEQ ID NO 285.  
 DE  
 XX Rabbit; probe; ss; array element; Parkinson's disease;  
 KW signalling pathway population; cancer; adenocarcinoma; leukaemia;  
 KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.  
 XX Oryctolagus cuniculus.  
 XX US6500938-B1.  
 PN 31-DEC-2002.  
 XX 98US-00016434.  
 PD 30-JAN-1998;  
 PF 98US-00016434.  
 XX 30-JAN-1998;  
 PR 98US-00016434.  
 XX  
 (INCY-) INCYTE GENOMICS INC.  
 PA Au-Young J, Seilhamer JJ;  
 XX WPI; 2003-352189/33.  
 DR Combination of polynucleotide probes, useful as array elements in a  
 PT microarray for monitoring the expression of a number of target  
 PT polynucleotides.  
 XX Claim 1; SEQ ID NO 285; 65pp; English.  
 PS The invention relates to a combination which, comprises a number of  
 CC polynucleotide probes comprising a sequence selected from one of the 1490  
 CC sequences mentioned in the specification. The combination is useful as an  
 CC array element in a microarray for monitoring the expression of a number  
 CC of target polynucleotides. The microarray is particularly useful in the  
 CC diagnosis and treatment of cancer and immunopathology and neuropathology.  
 CC The microarray is useful in diagnostics and treatment regimens, drug  
 CC discovery and development, toxicological and carcinogenicity studies,  
 CC forensics and pharmacogenomics. The microarray is also useful for  
 CC monitoring progression of diseases and for developing sophisticated  
 CC profiles for the effects of currently available therapeutic drugs. The  
 CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs  
 CC and genomic fragments and in research and diagnostic applications. The  
 CC array can detect changes in expression in a large number of genes coding  
 CC for different signalling pathway populations which can be used to diagnose  
 CC various diseases including cancer e.g. adenocarcinoma and leukaemia,  
 CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease  
 CC and Parkinson's disease. Note: The present sequence represents a polynucleotide  
 CC probe of the invention. The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=06500938B1  
 XX Sequence 202 BP; 36 A; 67 C; 56 G; 40 T; 0 U; 3 Other;  
 SQ  
 Alignment Scores: Length: 202  
 Pred. No.: 2.03e+03 Matches: 4  
 Score: 22.00 Conservative: 0  
 Percent Similarity: 80.00% Mismatches: 1  
 Best Local Similarity: 80.00% Indels: 0  
 Query Match: 91.67% Gaps: 0  
 DB: 10  
 US-10-030-194A-6 (1-6) x ACAS5687 (1-202)  
 QY 1 GlyTyr\*\*\*ValGlu 5  
 DB 86 GGCTATACAGTGGAG 72  
 RESULT 46  
 ADI55483/C  
 ID ADI55483 standard; DNA; 202 BP.  
 XX ADI55483;  
 AC 22-APR-2004 (first entry)  
 XX Human polynucleotide probe #285.  
 DE  
 XX Human; probe; ss; receptor-like polypeptide; transducing polypeptide;  
 KW effector-like polypeptide; cancer; immunopathology; neuropathology;  
 KW drug development; toxicology; carcinogenicity;  
 KW signalling pathway polypeptide; adrenal gland; bladder; bone;  
 KW bone marrow; brain; breast; cervix; tumour; immunopathology; AIDS;  
 KW diabetes; pancreatitis; osteoporosis; ulcerative colitis; neuropathology;  
 KW dementia; amnesia; epilepsy; Alzheimer's disease; depression.  
 XX Homo sapiens.  
 OS  
 XX US2004010136-A1.  
 PN  
 XX

PD 15-JAN-2004.  
XX  
PF 26-NOV-2002; 2002US-00305720.  
XX  
PR 30-JAN-1998; 98US-00016434.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Au-Young J, Seilhamer JJ;  
PI WPI; 2004-090520/09.  
XX  
XX New composition comprising polynucleotide probes, useful as array  
PT elements in a microarray for monitoring the expression of target  
PT polynucleotides or purifying a subpopulation of mRNAs, cDNA, or genomic  
PT fragments.  
XX  
XX Claim 6; SEQ ID NO 285; 73pp; English.  
XX  
XX The invention relates to a composition of polynucleotide probes  
CC comprising first polynucleotide probes comprising at least a portion of a  
CC gene encoding a receptor-like polypeptide, second polynucleotide probes  
CC comprising at least a portion of a gene encoding a transducing  
CC polypeptide and third polynucleotide probes comprising at least a portion  
CC of a gene encoding an effector-like polypeptide. The probes of the  
CC composition are useful as array elements in a microarray for monitoring  
CC the expression of target polynucleotides. The microarray is useful in the  
CC diagnosis and treatment of cancer, an immunopathology or a  
CC neuropathology. It can also be used for drug discovery and development,  
CC toxicological and carcinogenicity studies, forensics or pharmacogenomics.  
CC Microarrays can also be used for monitoring the progression of diseases  
CC that may be associated with the altered expression of signalling pathway  
CC polypeptides. The composition can also be used to purify a subpopulation  
CC of mRNAs, cDNAs, or genomic fragments in a sample. The expression profile  
CC is also useful for the diagnosis and treatment of cancer, e.g. cancers of  
CC the adrenal gland, bladder, bone, bone marrow, brain, breast or cervix,  
CC an immunopathology, e.g. AIDS, diabetes, pancreatitis, osteoporosis or  
CC ulcerative colitis, or a neuropathology, e.g. dementia, amnesia,  
CC epilepsy, Alzheimer's disease or depression. This sequence represents a  
CC human polynucleotide probe of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 202 BP; 36 A; 67 C; 56 G; 40 T; 0 U; 3 Other;

Alignment Scores:  
Pred. No.: 2.03e+03 Length: 202  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 12 Gaps: 0

US-10-030-194A-6 (1-6) x ADI55483 (1-202)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 86 GGCATACAGTGGAG 72

RESULT 47  
AAC31424  
ID AAC31424 standard; cDNA; 206 BP.  
XX  
XX AAC31424;  
XX  
XX 06-OCT-2000 (first entry)  
XX Human secreted protein 5' EST, SEQ ID NO: 35499.  
XX  
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX

OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-00200610.  
XX  
PR 26-FEB-1999; 99US-0122487P.  
XX  
XX (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
DR  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
PS Claim 1; SEQ ID NO 35499; 71pp + Sequence Listing; English.  
XX  
XX The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors  
XX  
SQ Sequence 206 BP; 57 A; 46 C; 32 G; 71 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.07e+03 Length: 206  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x AAC31424 (1-206)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 154 GGGTATGCTGTGAG 168

RESULT 48  
ADH84632/C  
ID ADH84632 standard; DNA; 225 BP.  
XX  
XX AC ADH84632;  
XX  
XX DT 22-APR-2004 (first entry)  
XX  
XX DE Enterococcus faecalis polynucleotide #2517.  
XX  
XX KW Enterococcus faecalis infection; transcription regulatory element;  
XX antibacterial; gene; ds.  
XX  
XX OS Enterococcus faecalis.  
XX  
XX PN US6617156-B1.  
XX  
XX PD 09-SEP-2003.  
XX  
PF 13-AUG-1998; 98US-00134000.

XX PR 15-AUG-1997; 9TUS-0055778P.  
 XX PA (DOUC/) DOUCETTE-STAMM L A.  
 XX PA (BUSH/) BUSH D.  
 XX PI Doucette-Stamm LA, Bush D;  
 XX XX WPI; 2003-895394/82.  
 XX DR P-P5DB; ADH88037.  
 XX PT New nucleic acid comprising a sequence encoding an Enterococcus faecalis  
 XX PT polypeptide, useful for preparing a composition for diagnosing or  
 XX PT treating E. faecalis infection.  
 XX PS Disclosure; SEQ ID NO 2517; 193pp; English.  
 XX CC The invention relates to Enterococcus faecalis polynucleotides and  
 XX CC polypeptides. The invention also relates to a recombinant expression  
 XX CC vector comprising a polynucleotide operably linked to a transcription  
 XX CC regulatory element, a cell comprising a recombinant vector, a method for  
 XX CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising  
 XX CC a sequence not given in the specification, a recombinant vector  
 XX CC comprising the nucleic acid and a cell comprising the recombinant vector.  
 XX CC The polynucleotides can be used to detect the presence of E. faecalis in  
 XX CC a sample. The sequences are useful for preparing a composition for  
 XX CC diagnosing or treating Enterococcus faecalis infection. This sequence  
 XX CC represents an E. faecalis polynucleotide of the invention.  
 XX SQ Sequence 225 BP; 94 A; 38 C; 37 G; 56 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.28e+03 Length: 225  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-030-194A-6 (1-6) x ADH84632 (1-225)  
 QY 1 GLYTyr\*\*\*ValGlu 5  
 Db 118 GGTTATAGTCTGCA 104  
 RESULT 49  
 ABX51080/C  
 ID ABX51080 standard; cDNA; 227 BP.  
 XX AC ABX51080;  
 XX DT 25-FEB-2003 (first entry)  
 XX DE Bovine EST associated with lactation/muscle/fat deposition #1009.  
 XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
 XX KW muscle deposition; fat deposition; genome mapping; gene identification;  
 XX KW gene analysis; cattle breeding.  
 XX OS Bos Taurus.  
 XX PN US2002137160-A1.  
 XX PD 26-SEP-2002.  
 XX PF 26-OCT-2001; 2001US-00983965.  
 XX PR 17-DEC-1998; 98US-0113678P.  
 XX PR 15-DEC-1999; 99US-00465231.  
 XX XX (BYAT/) BYATT J C.  
 XX PA (MATH/) MATHIALAGAN N.  
 XX PA (TAON/) TAO N.

PA (WARR/) WARREN W C.  
 XX BYATT JC, Mathialagan N, Tao N, Warren WC;  
 XX WPI; 2003-102386/09.  
 XX PT Purified nucleic acid molecules, useful for genome mapping, gene  
 XX PT identification and analysis, cattle breeding or preparation of constructs  
 XX PT for cattle gene expression and genetically improved cattle.  
 XX PS Claim 2; SEQ ID NO 1009; 38pp; English.  
 XX CC The invention relates to a purified nucleic acid molecule associated with  
 XX CC lactation or muscle and fat deposition (designated LMFD), derived from  
 XX CC cattle, and the LMFD nucleic acid can specifically hybridize to a second  
 XX CC nucleic acid molecule comprising any of 5912 nucleotide sequences,  
 XX CC appearing as ABX50072-ABX55983, or complements of them. Also included are  
 XX CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic  
 XX CC acid linked to a promoter and a 3' non-translated sequence that  
 XX CC functions in the cell to cause termination of transcription and addition  
 XX CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
 XX CC (2) determining a level or pattern of a molecule in a bovine cell or  
 XX CC tissue comprising: (a) incubating a marker nucleic acid (comprising any  
 XX CC of the 5912 nucleic acid sequences or its complement or fragment) with a  
 XX CC complementary nucleic acid molecule obtained from the bovine cell or  
 XX CC tissue, where hybridisation between the marker nucleic acid and the  
 XX CC complementary nucleic acid permits the detection of the molecule; and (b)  
 XX CC detecting the level or pattern of the complementary nucleic acid, where  
 XX CC the detection of the complementary nucleic acid is predictive of the  
 XX CC level or pattern of the molecule. The LMFD nucleic acid is used for  
 XX CC determining a level or pattern of a molecule in a bovine cell or tissue.  
 XX CC It is useful for genome mapping, gene identification and analysis, cattle  
 XX CC breeding, preparation of constructs for use in cattle gene expression, or  
 XX CC for genetically improving cattle. The present sequence is one of the 5912  
 XX CC bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present  
 XX CC sequence was not shown in the specification but was obtained in  
 XX CC electronic format from the USPTO web site:  
 XX CC seqdata.uspto.gov/sequence.html?DocID=20020137160  
 XX SQ Sequence 227 BP; 41 A; 68 C; 46 G; 72 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.3e+03 Length: 227  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 8 Gaps: 0  
 US-10-030-194A-6 (1-6) x ABX51080 (1-227)  
 QY 1 GLYTyr\*\*\*ValGlu 5  
 Db 213 GGATATACAGTGGAG 199  
 RESULT 50  
 ADH83521  
 ID ADH83521 standard; DNA; 228 BP.  
 XX AC ADH83521;  
 XX DT 22-APR-2004 (first entry)  
 XX DE Enterococcus faecalis polynucleotide #1406.  
 XX KW Enterococcus faecalis infection; transcription regulatory element;  
 XX KW antibacterial; gene; ds.  
 XX OS Enterococcus faecalis.  
 XX PN US6617156-B1.  
 XX XX 09-SEP-2003.  
 PD



XX PF 13-AUG-1998; 98US-00134000.  
 XX PR 15-AUG-1997; 97US-0055778P.  
 XX PA (DOUC/) DOUCETTE-STAMM L A.  
 XX PA (BUSH/) BUSH D.  
 XX PI Doucette-Stamm LA, Bush D;  
 XX WPI; 2003-895394/82.  
 XX DR P-PSDB; ADH86926.  
 XX PS New nucleic acid comprising a sequence encoding an Enterococcus faecalis  
 PT polypeptide, useful for preparing a composition for diagnosing or  
 PT treating E. faecalis infection.  
 XX PS Disclosure; SEQ ID NO 1406; 193pp; English.  
 XX CC The invention relates to Enterococcus faecalis polynucleotides and  
 CC polypeptides. The invention also relates to a recombinant expression  
 CC vector comprising a polynucleotide operably linked to a transcription  
 CC regulatory element, a cell comprising a recombinant vector, a method for  
 CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising  
 CC a sequence not given in the specification, a recombinant vector  
 CC comprising the nucleic acid and a cell comprising the recombinant vector.  
 CC The polynucleotides can be used to detect the presence of E. faecalis in  
 CC a sample. The sequences are useful for preparing a composition for  
 CC diagnosing or treating Enterococcus faecalis infection. This sequence  
 CC represents an E. faecalis polynucleotide of the invention.  
 XX SQ Sequence 228 BP; 88 A; 33 C; 41 G; 66 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.31e+03 Length: 228  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-030-194A-6 (1-6) x ADH83521 (1-228)  
 Qy 1 GlyTyr\*\*\*ValGlu 5  
 Db 131 GGCTATTTCAGTGGAA 145  
 RESULT 51  
 ABX41213  
 ID ABX41213 standard; cDNA; 241 BP.  
 XX AC ABX41213;  
 XX DT 20-FEB-2003 (first entry)  
 XX DE Bovine EST associated with lactation/muscle/fat deposition #6378.  
 XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
 KW muscle deposition; fat deposition; genome mapping; gene identification;  
 KW gene analysis; cattle breeding.  
 XX OS Bos Taurus.  
 XX PN US2002137139-A1.  
 XX PD 26-SEP-2002.  
 XX PF 24-SEP-2001; 2001US-00960352.  
 XX PR 12-JAN-1999; 98US-0115707P.  
 PR 11-JAN-2000; 2000US-00480902.  
 XX PA (BYATT/) BYATT J C.

PA (MATH/) MATHIALAGAN N.  
 PA (TAON/) TAO N.  
 XX (WARR/) WARREN W C.  
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
 XX WPI; 2003-110599/10.  
 XX PT New nucleic acid associated with lactation, and muscle and fat  
 PT deposition, useful for genome mapping, gene identification and analysis,  
 PT cattle breeding, or for genetically improving cattle.  
 XX PS Claim 2; SEQ ID NO 6378; 245pp; English.  
 XX CC The invention relates to a purified nucleic acid molecule associated with  
 CC lactation or muscle and fat deposition (designated LMFD), derived from  
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second  
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,  
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are  
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic  
 CC acid linked to a promoter and a 3' non-translated sequence that  
 CC functions in the cell to cause termination of transcription and addition  
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
 CC (2) determining a level or pattern of a molecule in a bovine cell or  
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any  
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a  
 CC complementary nucleic acid molecule obtained from the bovine cell or  
 CC tissue, where hybridisation between the marker nucleic acid and the  
 CC complementary nucleic acid permits the detection of the molecule; and (b)  
 CC detecting the level or pattern of the complementary nucleic acid, where  
 CC the detection of the complementary nucleic acid is predictive of the  
 CC level or pattern of the molecule. The LMFD nucleic acid is used for  
 CC determining a level or pattern of a molecule in a bovine cell or tissue.  
 CC It is useful for genome mapping, gene identification and analysis, cattle  
 CC breeding, preparation of constructs for use in cattle gene expression, or  
 CC for genetically improving cattle. The present sequence is one of the  
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The  
 CC present sequence was not shown in the specification but was obtained in  
 CC electronic format from the USPTO web site:  
 CC seqdata.uspto.gov/sequence.html?docID=20020137139  
 XX SQ Sequence 241 BP; 79 A; 44 C; 48 G; 70 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.46e+03 Length: 241  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 8 Gaps: 0  
 US-10-030-194A-6 (1-6) x ABX41213 (1-241)  
 Qy 1 GlyTyr\*\*\*ValGlu 5  
 Db 123 GGATATGCAGTAGAG 137  
 RESULT 52  
 ADA01774/c  
 ID ADA01774 standard; DNA; 247 BP.  
 XX AC ADA01774;  
 XX DT 06-NOV-2003 (first entry)  
 XX DE Mouse carcinoma associated nucleic acid, SEQ ID NO:293.  
 XX KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;  
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
 XX ds.  
 XX OS Mus sp.  
 XX

PN WO2003057146-A2.  
XX 17-JUL-2003.  
XX 26-DEC-2002; 2002WO-US041141.  
XX 26-DEC-2001; 2001US-00035832.  
XX (SAGR-) SAGRES DISCOVERY.  
XX Morris DW;  
XX WPI; 2003-587068/55.  
XX New recombinant nucleic acid encoding carcinoma associated protein,  
PT useful for preparing compositions for treating carcinomas.  
XX Claim 1; Page 158; 245pp; English.  
XX The invention relates to recombinant carcinoma associated (CA) nucleic  
CC acid sequences from mouse and human (ADA01482-ADA03094), and to  
CC recombinant carcinoma associated proteins (CAP) encoded by them. The  
CC invention also encompasses expression vectors and host cells comprising a  
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically  
CC binds to the protein, and a biochip comprising CA nucleic acid or  
CC fragments thereof. The sequences of the invention were identified using  
CC oncogenic retroviruses, which insert into the genome of the host organism  
CC at random. Many of these do not carry transduced host oncogenes or  
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a  
CC direct consequence of the effects of proviral integration into host  
CC genomes. The CA nucleic acid sequences can be used to diagnose  
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose  
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or  
CC leukaemia) or a propensity to carcinoma by determination of the sequence  
CC of a CA gene, or by determination of CA gene expression in particular  
CC tissues. CA nucleic acids, proteins and antibodies are also useful as  
CC therapeutic agents and in screening and evaluating drug candidates. The  
CC present sequence represents a specifically claimed murine CA nucleic acid  
CC sequence of the invention. Note: The sequence data for this patent is  
CC also available in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 247 BP; 54 A; 82 C; 57 G; 54 T; 0 U; 0 Other;  
Alignment Scores: Length: 247  
Pred. No.: 2.52e+03 Matches: 4  
Score: 22.00  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
Gaps: 0  
DB: 9  
US-10-030-194A-6 (1-6) x ADA01774 (1-247)  
QY 1 GlyTyr\*\*\*ValGlu 5  
DB 194 GGGTACGCGGTAGAG 180  
RESULT 53  
ADB71513/c  
ID ADB71513 standard; DNA; 247 BP.  
XX ADB71513;  
XX 04-DEC-2003 (first entry)  
XX Mouse carcinoma associated gene fragment #293.  
XX mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;  
XX cancer; neoplasm; adenocarcinoma; sarcoma.  
XX Mus sp.  
XX WO2003008583-A2.  
PN

XX 30-JAN-2003.  
XX 26-DEC-2001; 2001WO-US051291.  
XX 02-MAR-2001; 2001US-00798586.  
XX 23-OCT-2001; 2001US-00004113.  
XX 08-NOV-2001; 2001US-00052482.  
XX 30-NOV-2001; 2001US-00997722.  
XX 20-DEC-2001; 2001US-00034650.  
XX (SAGR-) SAGRES DISCOVERY.  
XX Morris DW, Engelhard EK;  
XX WPI; 2003-239337/23.  
XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,  
PT cancers, neoplasm, adenocarcinoma, or sarcomas.  
XX Claim 1; Page 124; 2304pp; English.  
XX The invention relates to a novel recombinant nucleic acid comprising a  
CC nucleotide sequence selected from any of the 660 sequences fully defined  
CC in the specification. A polynucleotide of the invention has cytostatic  
CC activity, and may have a use in gene therapy, or in a vaccine. The  
CC recombinant nucleic acids and polypeptides are useful for treating  
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and  
CC sarcomas. The sequences shown in ADB71221-ADB72172 represent mouse  
CC sequence tags, or genomic insertion sites, of carcinoma associated (CA)  
CC genes of the invention.  
XX SQ Sequence 247 BP; 54 A; 82 C; 57 G; 54 T; 0 U; 0 Other;  
Alignment Scores: Length: 247  
Pred. No.: 2.52e+03 Matches: 4  
Score: 22.00  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
Gaps: 0  
DB: 10  
US-10-030-194A-6 (1-6) x ADB71513 (1-247)  
QY 1 GlyTyr\*\*\*ValGlu 5  
DB 194 GGGTACGCGGTAGAG 180  
RESULT 54  
AAH81246/c  
ID AAH81246 standard; DNA; 251 BP.  
XX AAH81246;  
XX 21-SEP-2001 (first entry)  
XX Escherichia coli nucleotide sequence SEQ ID NO:45.  
XX Escherichia coli; identification; proliferation; microorganism;  
KW antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;  
KW bacterial growth inhibition; ds.  
XX Escherichia coli.  
XX WO200148209-A2.  
XX 05-JUL-2001.  
XX 19-DEC-2000; 2000WO-US034419.  
XX 23-DEC-1999; 99US-0173005P.  
XX (ELIT-) ELITRA PHARM INC.  
PA

XX Forsyth RA, Ohlsen KL, Zyskind JW;  
 XX WPI; 2001-457376/49.  
 XX Novel nucleic acids encoding proteins required for Escherichia coli  
 PT proliferation, useful for screening for antimicrobial agents.  
 XX Claim 1; Page 120; 596pp; English.  
 XX The present invention describes a purified or isolated nucleic acid  
 CC sequence (I) consisting essentially of one of the 93 nucleotide sequences  
 CC given in AAH81202 to AAH81294, where expression of the nucleic acid in a  
 CC microorganism is capable of inhibiting proliferation of a microorganism.  
 CC (I) have antibacterial and antibiotic activities, and can be used in gene  
 CC therapy. Expression of (I) in a microorganism inhibits proliferation of  
 CC the microorganism, and the manufactured antibiotic is useful for reducing  
 CC the activity or level of a gene product required for proliferation of a  
 CC microorganism in a subject, specifically humans. The nucleic acids that  
 CC inhibit bacterial growth or proliferation can be used as antisense  
 CC therapeutics for killing bacteria. In addition to therapeutic  
 CC applications, the nucleic acid sequences complementary to sequences  
 CC required for proliferation can be used as diagnostic tools. For example,  
 CC nucleic acid probes complementary to proliferation-required sequences  
 CC that are specific for particular species of microorganisms can be used as  
 CC probes to identify particular microorganism species in clinical  
 CC specimens. AAH81295 to AAH81487 encode the Escherichia coli proteins  
 CC given in AAG98239 to AAG98431, and AAH81488 to AAH81491 represent  
 CC oligonucleotides, which are used in the exemplification of the present  
 CC invention  
 XX SQ Sequence 251 BP; 67 A; 62 C; 60 G; 62 T; 0 U; 0 Other;  
 Alignment Scores: Length: 251  
 Pred. No.: 2.57e+03 Matches: 4  
 Score: 22.00 Conservative: 0  
 Percent Similarity: 80.00% Mismatches: 1  
 Best Local Similarity: 80.00% Indels: 0  
 Query Match: 91.67% Gaps: 0  
 DB: 5  
 US-10-030-194A-6 (1-6) x AAH81246 (1-251)  
 Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 242 GGTACACCGTCGAA 228  
 RESULT 55  
 ACA13575/c  
 ID ACA13575 standard; DNA; 251 BP.  
 XX ACA13575;  
 AC ACA13575;  
 XX 27-OCT-2003 (revised)  
 DT 19-JUN-2003 (first entry)  
 XX Query Match: 91.67%  
 DE Prokaryotic essential gene antisense oligonucleotide #1445.  
 XX Antisense; ss; prokaryotic essential gene; cell proliferation;  
 KW drug design.  
 XX Archaea.  
 OS  
 FN WO200277183-A2.  
 XX 03-OCT-2002.  
 PD  
 XX 21-MAR-2002; 2002WO-US009107.  
 PF  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX Claim 1; SEQ ID NO 1445; 1766pp; English.  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the 6213  
 CC antisense sequences of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 27-OCT-2003 to  
 CC standardise OS field)  
 XX SQ Sequence 251 BP; 67 A; 62 C; 61 G; 61 T; 0 U; 0 Other;  
 Alignment Scores: Length: 251  
 Pred. No.: 2.57e+03 Matches: 4  
 Score: 22.00 Conservative: 0  
 Percent Similarity: 80.00% Mismatches: 1  
 Best Local Similarity: 80.00% Indels: 0  
 Query Match: 91.67% Gaps: 0  
 DB: 8  
 US-10-030-194A-6 (1-6) x ACA13575 (1-251)  
 Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 242 GGTACACCGTCGAA 228  
 RESULT 56  
 AAK89055  
 ID AAK89055 standard; DNA; 265 BP.  
 XX AAK89055;  
 AC AAK89055;  
 XX 05-NOV-2001 (first entry)  
 DT  
 XX Human digestive system antigen genomic sequence SEQ ID NO: 2631.  
 DE



PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Barash SC, Ruben SM;  
XX DR WPI; 2001-502630/55.  
XX PT Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases.  
XX PS Disclosure: SEQ ID NO 2631; 986pp; English.  
XX CC The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a genomic DNA fragment  
CC encoding a digestive system antigen of the invention  
XX SQ Sequence 265 BP; 87 A; 32 C; 54 G; 92 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 2,72e+03 Length: 265  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0  
  
US-10-030-194a-6 (1-6) x AAK89055 (1-265)  
  
Qy 1 GlyTyr\*\*\*Valglu 5  
Db 52 GTTATTCA GTGAG 66  
  
RESULT 57  
AAS39651  
ID AAS39651 standard; DNA; 265 BP.  
XX AC AAS39651;  
XX DT 17-DEC-2001 (first entry)  
XX DE Genomic sequence #70 encoding human colon associated polypeptide.  
XX KW Human; colon cancer; congenital abnormality; infection; colitis;  
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;  
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;  
KW sigmoid disease; antibacterial; antiviral; antiinflammatory; cytostatic;  
KW ds.  
XX OS Homo sapiens.  
XX FN WO200155302-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US001240.  
XX PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231988P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.

20-OCT-2000; 2000US-0241786P.  
20-OCT-2000; 2000US-0241787P.  
20-OCT-2000; 2000US-0241808P.  
20-OCT-2000; 2000US-0241809P.  
20-OCT-2000; 2000US-0241826P.  
01-NOV-2000; 2000US-0244617P.  
08-NOV-2000; 2000US-0244617P.  
08-NOV-2000; 2000US-0246475P.  
08-NOV-2000; 2000US-0246476P.  
08-NOV-2000; 2000US-0246477P.  
08-NOV-2000; 2000US-0246478P.  
08-NOV-2000; 2000US-0246523P.  
08-NOV-2000; 2000US-0246524P.  
08-NOV-2000; 2000US-0246525P.  
08-NOV-2000; 2000US-0246526P.  
08-NOV-2000; 2000US-0246527P.  
08-NOV-2000; 2000US-0246528P.  
08-NOV-2000; 2000US-0246532P.  
08-NOV-2000; 2000US-0246609P.  
08-NOV-2000; 2000US-0246610P.  
08-NOV-2000; 2000US-0246611P.  
08-NOV-2000; 2000US-0246613P.  
17-NOV-2000; 2000US-0249207P.  
17-NOV-2000; 2000US-0249208P.  
17-NOV-2000; 2000US-0249209P.  
17-NOV-2000; 2000US-0249210P.  
17-NOV-2000; 2000US-0249211P.  
17-NOV-2000; 2000US-0249212P.  
17-NOV-2000; 2000US-0249213P.  
17-NOV-2000; 2000US-0249214P.  
17-NOV-2000; 2000US-0249215P.  
17-NOV-2000; 2000US-0249216P.  
17-NOV-2000; 2000US-0249217P.  
17-NOV-2000; 2000US-0249218P.  
17-NOV-2000; 2000US-0249244P.  
17-NOV-2000; 2000US-0249245P.  
17-NOV-2000; 2000US-0249264P.  
17-NOV-2000; 2000US-0249265P.  
17-NOV-2000; 2000US-0249297P.  
17-NOV-2000; 2000US-0249299P.  
01-DEC-2000; 2000US-0249300P.  
01-DEC-2000; 2000US-0250160P.  
01-DEC-2000; 2000US-0250391P.  
03-DEC-2000; 2000US-0251030P.  
03-DEC-2000; 2000US-0251988P.  
05-DEC-2000; 2000US-0256719P.  
05-DEC-2000; 2000US-0256719P.  
08-DEC-2000; 2000US-0251856P.  
08-DEC-2000; 2000US-0251868P.  
08-DEC-2000; 2000US-0251869P.  
08-DEC-2000; 2000US-0251989P.  
08-DEC-2000; 2000US-0251990P.  
11-DEC-2000; 2000US-0254097P.  
05-JAN-2001; 2001US-02559678P.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-465567/50.  
Isolated polypeptide for treating, preventing and/ or prognosing  
disorders related to the colon including colon cancers and also for  
testing and detection e.g. diagnosis.  
Disclosure; SEQ ID NO 548; 562pp; English.  
The present invention relates to the isolation of novel human colon  
associated polypeptides (AAU22701), and the cDNA and genomic  
sequences encoding for them. The sequences of the invention are useful in  
the diagnosis, treatment, prevention and/or prognosis of disorders of the  
colon including colon cancer, congenital abnormalities (e.g. atresia and  
stenosis), bacterial and viral infections, inflammatory bowel disease

CC (IBD), neoplastic cell disorders (e.g. polyps and adenomas, intestinal  
CC inflammatory disorders, colitis, colonic inflammation, diarrhoea and  
CC dysentery, malabsorption syndromes (e.g. lactose intolerance), intestinal  
CC obstruction and sigmoid diseases. The polynucleotides sequences of the  
CC invention can also be used in gene therapy. AAS39582-AAS40060 represent  
CC DNA sequences encoding for the novel human colon associated polypeptides  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 265 BP; 87 A; 32 C; 54 G; 92 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 2.72e+03 Length: 265  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 5 Gaps: 0  
US-10-030-194A-6 (1-6) x AAS39651 (1-265)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 52 GGTATTTCAGTGAG 66  
RESULT 58  
ADB32611  
ID ADB32611 standard; DNA; 265 BP.  
XX ADB32611;  
XX 04-DEC-2003 (first entry)  
XX Human novel colon related polypeptide DNA SEQ ID NO 548.  
XX gene therapy; cancer; liver disorder; hepatitis; neural disorder;  
XX Alzheimer's disease; human; colon; ds.  
XX Homo sapiens.  
OS  
US2003050231-A1.  
XX  
PD 13-MAR-2003.  
XX  
PF 17-JAN-2001; 2001US-00764872.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205151P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.

PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-02266681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226686P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0251990P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	26-SEP-2000;	2000US-0234998P.	XX	(ROSE/) ROSEN C A.	
PR	26-SEP-2000;	2000US-0235484P.	PA	(RUBE/) RUBEN S M.	
PR	27-SEP-2000;	2000US-0235834P.	PA	(BARA/) BARASH S C.	
PR	27-SEP-2000;	2000US-0235836P.	XX	Rosen CA, Ruben SM, Barash SC;	
PR	29-SEP-2000;	2000US-0236327P.	PI		
PR	29-SEP-2000;	2000US-0236367P.	XX		
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	DR	WPI; 2003-625420/59.	
PR	29-SEP-2000;	2000US-0236370P.	XX		
PR	02-OCT-2000;	2000US-0236802P.	PT	New nucleic acid molecule, useful for preparing a medicament for	
PR	02-OCT-2000;	2000US-0237037P.	PT	preventing, treating or ameliorating a medical condition e.g. cancer,	
PR	02-OCT-2000;	2000US-0237038P.	PT	liver disorders or neural disorders.	
PR	02-OCT-2000;	2000US-0237039P.	XX		
PR	02-OCT-2000;	2000US-0237040P.	FS	Disclosure; SEQ ID NO 548; 216pp; English.	
PR	13-OCT-2000;	2000US-0239935P.	XX		
PR	20-OCT-2000;	2000US-0239937P.	CC	The invention relates to an isolated nucleic acid molecule encoding a	
PR	20-OCT-2000;	2000US-0240960P.	CC	polypeptide. The nucleic acid is useful for preparing a medicament for	
PR	20-OCT-2000;	2000US-0241221P.	CC	preventing, treating or ameliorating a medical condition e.g. cancer,	
PR	20-OCT-2000;	2000US-0241785P.	CC	liver disorders such as hepatitis or neural disorders such as Alzheimer's	
PR	20-OCT-2000;	2000US-0241786P.	CC	disease. The present sequence represents a human novel colon related	
PR	20-OCT-2000;	2000US-0241787P.	CC	polypeptide DNA. Note: The sequence data for this patent did not form	
PR	20-OCT-2000;	2000US-0241808P.	CC	part of the printed specification but was obtained in electronic format	
PR	20-OCT-2000;	2000US-0241809P.	CC	directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030050231.	
PR	01-NOV-2000;	2000US-0241826P.	XX		
PR	01-NOV-2000;	2000US-0244617P.	SQ	Sequence 265 BP; 87 A; 32 C; 54 G; 92 T; 0 U; 0 Other;	
PR	08-NOV-2000;	2000US-0246475P.			
PR	08-NOV-2000;	2000US-0246476P.			
PR	08-NOV-2000;	2000US-0246477P.			
PR	08-NOV-2000;	2000US-0246478P.			
PR	08-NOV-2000;	2000US-0246523P.			
PR	08-NOV-2000;	2000US-0246524P.			
PR	08-NOV-2000;	2000US-0246525P.			
PR	08-NOV-2000;	2000US-0246526P.			
PR	08-NOV-2000;	2000US-0246527P.			
PR	08-NOV-2000;	2000US-0246528P.			
PR	08-NOV-2000;	2000US-0246529P.			
PR	08-NOV-2000;	2000US-0246528P.			

Alignment Scores:

Pred. No.:	2.72e+03	Length:	265
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	9	Gaps:	0

US-10-030-194A-6 (1-6) x ADB32611 (1-265)

Thu Nov 4 17:32:32 2004

Ov 1 GlyTyr\*\*\*ValGlu 5  
Db 52 GGTATTTCAGTGGAG 66  
RESULT 59  
ID AAK87739 standard; cDNA; 267 BP.  
AC AAK87739;  
XX AAK87739;  
DT 05-NOV-2001 (first entry)  
DE Human digestive system antigen coding sequence SEQ ID NO: 55.  
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;  
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
KW digestive system disorder; Meckel's diverticulum; ss.  
XX Homo sapiens.  
PN WO200155314-A2.  
XX 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US001324.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205151P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.



```

PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-502630/55.
XX P-PSDB; AAM91966.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases.
XX
XX Claim 1; SEQ ID NO 55; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a cDNA encoding a digestive
XX system antigen of the invention
XX
XX Sequence 267 BP; 86 A; 30 C; 56 G; 90 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 2,75e+03 Length: 267
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x AAK97739 (1-267)
QY 1 GlyTyr***ValGlu 5
Db 55 GGTATTTCAGTGGAG 69

RESULT 60
AAS39374
ID AAS39374 standard; cDNA; 267 BP.
XX
AC AAS39374;
XX
DT 17-DEC-2001 (first entry)
XX
XX cDNA encoding novel human colon associated polypeptide #27.
XX
XX Human; colon cancer; congenital abnormality; infection; colitis;
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;
KW sigmoid disease; antibacterial; antiviral; antiinflammatory; cytostatic;
KW ss.
XX
XX Homo sapiens.
OS
XX
XX WO200155302-A2.
PN

```

```

XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001240.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 25-SEP-2000; 2000US-0234998P.
XX 26-SEP-2000; 2000US-0235484P.
XX 27-SEP-2000; 2000US-0235834P.
XX 27-SEP-2000; 2000US-0235836P.
XX 29-SEP-2000; 2000US-0236327P.

```

29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239933P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241828P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 01-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254037P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
XX  
XX  
PI  
XX

WPI; 2001-465567/50.  
P-PSDB; AAU22494.  
Isolated polypeptide for treating, preventing and/ or prognosing  
disorders related to the colon including colon cancers and also for  
testing and detection e.g. diagnosis.  
XX  
PS Claim 4; SEQ ID NO 37; 562pp; English.  
XX  
XX The present invention relates to the isolation of novel human colon  
associated polypeptides (AAU22468-AAU22701), and the cDNA and genomic  
sequences encoding for them. The sequences of the invention are useful in  
the diagnosis, treatment, prevention and/or prognosis of disorders of the  
colon including colon cancer, congenital abnormalities (e.g. atresia and  
stenosis), bacterial and viral infections, inflammatory bowel disease  
(IBD), neoplastic cell disorders (e.g. polyps and adenomas, intestinal  
inflammatory disorders, colitis, colonic inflammation, diarrhoea and  
dysentery, malabsorption syndromes (e.g. lactose intolerance), intestinal  
obstruction and sigmoid diseases. The polynucleotide sequences of the  
invention can also be used in gene therapy. AAS39348-AAS39581 represent  
cDNA sequences encoding for the novel human colon associated polypeptides  
of the invention. Note: The sequence data for this patent did not form  
part of the printed specification, but was obtained in electronic format  
directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 267 BP; 86 A; 30 C; 56 G; 90 T; 0 U; 5 Other;

Alignment Scores: 2.75e+03 Length: 267  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 5

US-10-030-194A-6 (1-6) x AAS39374 (1-267)

Qy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 55 GGTATTTCAGTGGAG 69

RESULT 61  
ADB32100  
ID ADB32100 standard; cDNA; 267 BP.  
XX  
AC ADB32100;  
XX  
XX 04-DEC-2003 (first entry)  
XX Human novel colon related polypeptide cDNA SEQ ID NO 37.  
XX gene therapy; ss; gene; cancer; liver disorder; hepatitis;  
KW neural disorder; Alzheimer's disease; human; colon.  
XX Homo sapiens.  
XX OS  
XX US2003050231-A1.  
XX PD  
XX 13-MAR-2003.  
XX  
XX 17-JAN-2001; 2001US-00764872.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR

PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241788P.  
PR 20-OCT-2000; 2000US-0241789P.  
PR 20-OCT-2000; 2000US-0241790P.  
PR 20-OCT-2000; 2000US-0241791P.  
PR 20-OCT-2000; 2000US-0241792P.  
PR 20-OCT-2000; 2000US-0241793P.  
PR 20-OCT-2000; 2000US-0241794P.  
PR 20-OCT-2000; 2000US-0241795P.  
PR 20-OCT-2000; 2000US-0241796P.  
PR 20-OCT-2000; 2000US-0241797P.  
PR 20-OCT-2000; 2000US-0241798P.  
PR 20-OCT-2000; 2000US-0241799P.  
PR 20-OCT-2000; 2000US-0241800P.  
PR 20-OCT-2000; 2000US-0241801P.  
PR 20-OCT-2000; 2000US-0241802P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR (ROSE/) ROSEN C A.  
PR (RUBE/) RUBEN S M.  
PR (BARA/) BARASH S C.  
XX Rosen CA, Ruben SM, Barash SC;  
XX  
XX WPI; 2003-625420/59.  
XX P-PSDB; ADB32334.  
XX  
XX New nucleic acid molecule, useful for preparing a medicament for  
XX preventing, treating or ameliorating a medical condition e.g. cancer,  
XX liver disorders or neural disorders.  
XX  
XX Claim 3; SEQ ID NO 37; 216pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule encoding a  
XX polypeptide. The nucleic acid is useful for preparing a medicament for  
XX preventing, treating or ameliorating a medical condition e.g. cancer,  
XX liver disorders such as hepatitis or neural disorders such as Alzheimer's  
XX disease. The present sequence represents a human cDNA encoding a novel

Thu Nov 4 17:32:32 2004

us-10-030-194a-6.p2n.rng

CC colon related polypeptide. Note: The sequence data for this patent did  
CC not form part of the printed specification but was obtained in electronic  
CC format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030050231.

XX  
SQ Sequence 267 BP; 86 A; 30 C; 56 G; 90 T; 0 U; 5 Other;

Alignment Scores: Pred. No.: 2.75e+03 Length: 267  
Matches: 4  
Score: 22.00  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
Gaps: 0

US-10-030-194A-6 (1-6) x ADB32100 (1-267)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 55 GGTATTTCAGTGGAG 69

RESULT 62  
AAK58037  
ID AAK58037 standard; cDNA; 269 BP.

AC AAK58037;

DT 06-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:3037.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ss.

OS Homo sapiens.

XX WO200157182-A2.

PD 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-483426/52.  
XX P-PSDB; AAM85256.  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX Claim 1; SEQ ID NO 3097; 3071bp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK67694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 269 BP; 48 A; 71 C; 91 G; 53 T; 0 U; 6 Other;

Alignment Scores:  
Pred. No.: 2,77e+03 Length: 269  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x AAK58037 (1-269)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db ||||| |||||  
6 GGTATTCCGTCGAG 20  
RESULT 63  
AAH20035  
ID AAH20035 standard; cDNA; 273 BP.  
XX  
XX AAH20035;  
XX  
XX 06-AUG-2001 (first entry)  
XX  
XX Mouse differentially expressed transcript (DET) A SEQ ID NO:1.  
XX  
XX Mouse; differentially expressed transcript; DET; identification;  
XX cell transplantation; neural damage repair; differential display;  
XX gene expression; neural stem cell; damaged vertebrate brain; ss.  
XX  
XX Mus sp.  
XX WO200131058-A2.  
XX  
XX 03-MAY-2001.  
XX  
XX 24-OCT-2000; 2000WO-GB004103.  
XX  
XX 25-OCT-1999; 99GB-00025210.  
XX  
XX (RENE-) RENEURON LTD.  
XX  
XX Price J, Uwanogho D;  
XX  
XX WPI; 2001-308658/32.  
XX  
XX Selecting neural stem cells capable of aiding repair when transplanted  
XX into a damaged vertebrate brain by comparing their gene expression with  
XX that of control cells known to aid repair.  
XX  
XX Claim 6; Page 13; 15pp; English.  
XX  
XX The present invention describes a method for selecting cells suitable for  
XX transplantation into a damaged vertebrate brain comprising comparing the  
XX gene expression profile with a control cell known to be suitable for  
XX transplantation. The method comprises: (a) selecting cells that are, or  
XX are capable of differentiating into, neural cells; (b) obtaining the gene  
XX expression profile of the cells; (c) comparing that profile with one from  
XX a control cell known to be suitable for transplantation; and (d)  
XX selecting cells with a similar profile to the control. Also described  
XX are: (i) a method identifying a gene whose expression can determine if a  
XX neural cell can repair a damaged brain, comprising: (a) carrying out  
XX steps (a) to (c) of the above method; and (b) isolating genes that are  
XX the same (or different) from those expressed by the control; (ii)  
XX selecting a cell suitable for transplantation into a damaged brain,  
XX comprising determining the presence of the 273, 325, 272, 438 or 206  
XX nucleotide sequences given in AAH20035 to AAH20039. The method is used to  
XX select neural stem cells which can be used to aid repair of a damaged  
XX vertebrate brain. The present sequence represents a mouse differentially  
XX expressed transcript (DET), which is used in the exemplification of the  
XX present invention  
XX  
SQ Sequence 273 BP; 71 A; 66 C; 74 G; 62 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,81e+03 Length: 273  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 5 Gaps: 0

US-10-030-194A-6 (1-6) x AAH20035 (1-273)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 126 GGATATCCCGTGGAA 140

RESULT 64  
ABX14653  
ID ABX14653 standard; cDNA; 273 BP.  
XX  
AC ABX14653;  
XX  
DT 06-MAR-2003 (first entry)  
XX  
DE Mouse differential display cDNA #1.  
XX  
KW Mouse; ss; differential display; multipotent neural cell;  
KW neural cell transplantation; MHP36; gene expression profile;  
KW brain damage.  
XX  
OS Mus sp.  
XX  
PN US6465215-B1.  
XX  
PD 15-OCT-2002.  
XX  
PF 25-OCT-2000; 2000US-00696569.  
XX  
PR 14-DEC-1999; 99US-0170692P.  
XX  
PA (RENE-) RENEURON LTD.  
XX  
PI Price J, Uwanogho D;  
XX  
PD WPI; 2003-147065/14.  
XX  
PT Determining suitability of candidate cell for transplantation, by  
PT identifying and comparing gene expression profiles of the cell and  
PT several control cells, and identifying the suitable cell for  
PT transplantation.  
XX  
PS Claim 1; Col 7-8; 8pp; English.  
XX  
CC The invention relates to determining the suitability of a candidate cell  
CC for transplantation into a damaged vertebrate brain for repairing damage  
CC to the brain, or identifying a gene, comprising: (a) selecting a  
CC candidate cell with a neural cell phenotype (e.g. a multipotent neural  
CC cell such as MHP36), or which differentiates into a cell with a neural  
CC cell phenotype (b) obtaining the gene expression profile of the candidate  
CC cell; (c) identifying a gene expression profile common to several control  
CC cells known to be suitable for transplantation, where the gene expression  
CC profile common to several control cells is selected from: (i) at least  
CC one expressed gene selected from a the differential display cDNAs  
CC appearing as ABX14653-ABX14657; (ii) at least one non-expressed gene  
CC selected from ABX14658-ABX14659; or (iii) both (i) and (ii); (d)  
CC comparing the gene expression profile of the candidate cell with the gene  
CC expression profile common to several control cells; and (e) identifying  
CC the candidate cell as being suitable for transplantation, if the  
CC candidate cell shares the gene expression profile of several control  
CC cells, or identifying a gene within the gene expression profile of the  
CC candidate cell or within the gene expression profile common to the  
CC several control cells, where the gene is expressed within the candidate  
CC cell but is not expressed within the several control cells, or where the  
CC gene is not expressed within the candidate cell but is expressed within  
CC the several control cells. The method is used for determining the  
CC suitability of a candidate cell for transplantation into a damaged  
CC vertebrate brain for repairing damage to brain, or for identifying a  
CC gene, the expression of which determines whether of not a neural cell can  
CC repair a damaged brain, where the candidate cell is undifferentiated, or  
CC multipotent. The present sequence is a cDNA fragment from one of the  
CC differentially displayed transcripts expressed in MHP36  
XX  
XX Sequence 273 BP; 71 A; 66 C; 74 G; 62 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.81e+03 Length: 273  
Score: 22.00 Matches: 4  
Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 10

US-10-030-194A-6 (1-6) x ABX14653 (1-273)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 126 GGATATCCCGTGGAA 140

RESULT 65  
ABZ55211  
ID ABZ55211 standard; cDNA; 276 BP.  
XX  
AC ABZ55211;  
XX  
DT 28-MAR-2003 (first entry)  
XX  
DE Aspergillus oryzae polynucleotide SEQ ID NO 4324.  
XX  
KW Aspergillus oryzae; fermentation; fungus; industrial; EST;  
KW expressed sequence tag; gene; ss.  
XX  
OS Aspergillus oryzae.  
XX  
PN WO200279476-A1.  
XX  
PD 10-OCT-2002.  
XX  
PF 22-MAR-2002; 2002WO-IB000890.  
XX  
PR 30-MAR-2001; 2001JP-00098371.  
XX  
PA (NABD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
PA (NARE-) NAT RES INST BREWING.  
PA (NORQ) NAT FOOD RES INST MIN AGRIC.  
XX  
PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;  
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;  
XX  
XX WPI; 2003-046817/04.  
XX  
DR Detection of expression of specific Aspergillus genes for monitoring the  
XX fermentation and growth conditions of the fungus, using DNA probes.  
XX  
PT Claim 1; SEQ ID NO 4324; 48pp + Sequence Listing; Japanese.  
XX  
PS The invention relates to a polynucleotide having any of 6006 specific  
XX sequences (ABZ50888-ABZ56893), which are expressed by a fungus under  
XX specific culture conditions including one or more of eutrophic,  
XX oligotrophic, solid, early germination, alkaline, high temperature, low  
XX temperature or maltose culture or polynucleotides stringently hybridising  
XX to these sequences. The polynucleotides are useful for monitoring the  
XX progress of fermentation and the growth conditions of a fungus,  
XX especially of Aspergillus oryzae which is widely used in industrial  
XX fermentation. Also monitoring for fungal contamination. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 276 BP; 75 A; 53 C; 65 G; 82 T; 0 U; 1 Other;

Alignment Scores:  
Pred. No.: 2.85e+03 Length: 276  
Score: 22.00 Matches: 4  
Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 8

US-10-030-194A-6 (1-6) x ABZ55211 (1-276)

Oy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 64 GGATATGCGGTGAG 78

RESULT 66

AAC23928  
ID AAC23928 standard; cDNA; 277 BP.

XX AC AAC23928;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein 5' EST, SEQ ID NO: 28003.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX KW gene therapy; chromosome mapping; ss.

XX OS Homo sapiens.

XX FN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-00200610.

XX PR 26-FEB-1999; 99US-0122487P.

XX PA (GEST ) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.

XX FT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX FT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX PS Claim 1; SEQ ID NO 28003; 71pp + Sequence Listing; English.

XX CC The present sequence is one of a large number of 5' ESTs derived from  
XX CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
XX CC identified within the present sequence. The 5' ESTs were prepared from  
XX CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
XX CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
XX CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
XX CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
XX CC derived from the 5' ends of mRNAs and even in those cases where longer  
XX CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
XX CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
XX CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
XX CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX CC They are used to obtain upstream regulatory sequences and to design  
XX CC expression and secretion vectors

XX SQ Sequence 277 BP; 66 A; 80 C; 66 G; 63 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 2.86e+03 Length: 277  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x AAC23928 (1-277)

Oy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 105 GGCTACACAGTAGAA 119

RESULT 67

AAQ60834  
ID AAQ60834 standard; DNA; 281 BP.

XX AC AAQ60834;

XX DT 25-MAR-2003 (revised)

XX DT 16-MAR-1994 (first entry)

XX DE Human brain Expressed Sequence Tag EST00937.

XX KW Gene transcription product; genetic markers; tagging; in vivo;  
XX KW transcription; mapping; locations; chromosomes; chromosomal; ss.

XX OS Homo sapiens.

XX FN WO9316178-A2.

XX PD 19-AUG-1993.

XX PF 12-FEB-1993; 93WO-US001294.

XX PR 12-FEB-1992; 92US-00837195.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.

XX PI Venter CJ, Adams MD, Moreno RF;

XX DR WPI; 1993-272882/34.

XX PT Enriched oligonucleotides and corresp. sequences - used as markers for  
XX PT human genes transcribed in-vivo, facilitate tagging of most human genes.

XX PS Example 4; Page 398; 500pp; English.

XX CC The Expressed Sequence Tag was isolated from a human brain cDNA library  
XX CC as part of a large set of ESTs which can be used as markers for human  
XX CC genes transcribed in vivo. They can be used to facilitate tagging of most  
XX CC human genes, for mapping locations of expressed genes on chromosomes, for  
XX CC individual or forensic identification, for mapping locations of disease-  
XX CC associated genes, for identification of tissue type, and for prepn. of  
XX CC antisense sequences, probes and constructs. EST00937 has a "poor" coding  
XX CC probability as evaluated using the coding-region prediction program CRM.  
XX CC See also AAQ59041-061440. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 281 BP; 45 A; 81 C; 91 G; 61 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 2.9e+03 Length: 281  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x AAQ60834 (1-281)

Oy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 41 GGTTACAGTGTGGAG 55

RESULT 68

ABZ52598/c

ID ABZ52598 standard; cDNA; 282 BP.

XX AC ABZ52598;

XX DT 28-MAR-2003 (first entry)

XX DE Aspergillus oryzae polynucleotide SEQ ID NO 1711.

XX KW Aspergillus oryzae; fermentation; fungus; industrial; EST;  
XX KW expressed sequence tag; gene; ss.

```

XX OS Aspergillus oryzae.
XX PN WO200279476-A1.
XX PD 10-OCT-2002.
XX PF 22-MAR-2002; 2002WO-IB000890.
XX PR 30-MAR-2001; 2001JP-00098371.
XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX PA (NARE-) NAT RES INST BREWING.
XX PA (NORQ) NAT FOOD RES INST MIN AGRIC.
XX PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
XX PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
XX DR WPI; 2003-046817/04.
XX PT Detection of expression of specific Aspergillus genes for monitoring the
XX PT fermentation and growth conditions of the fungus, using DNA probes.
XX PS Claim 1; SEQ ID NO 1711; 49pp + Sequence Listing; Japanese.
XX CC The invention relates to a polynucleotide having any of 6006 specific
XX CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
XX CC specific culture conditions including one or more of eutrophic,
XX CC oligotrophic, solid, early germination, alkaline, high temperature, low
XX CC temperature or maltose culture or polynucleotides stringently hybridising
XX CC to these sequences. The polynucleotides are useful for monitoring the
XX CC progress of fermentation and the growth conditions of a fungus,
XX CC especially of Aspergillus oryzae which is widely used in industrial
XX CC fermentation. Also monitoring for fungal contamination. Note: the
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 282 BP; 82 A; 43 C; 84 G; 71 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 2,91e+03 Length: 282
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x ABZ52598 (1-282)
QY 1 GlyTyr***ValGlu 5
Db 81 GGGTACACTGTCGAG 67
RESULT 69
AAC22592/C
ID AAC22592 standard; cDNA; 285 BP.
XX AC AAC22592;
XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 26667.
XX KW Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN EP1033401-A2.
XX PF 06-SEP-2000.
XX

```

```

PF 21-FEB-2000; 2000EP-00200610.
XX PR 26-FEB-1999; 99US-0122487P.
XX PA (GEST) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX CC New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX CC obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX PS Claim 1; SEQ ID NO 26667; 71pp + Sequence Listing; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX CC identified within the present sequence. The 5' ESTs were prepared from
XX CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX CC sequences usually correspond mainly to the 3' untranslated region (UTR)
XX CC of the mRNA because they are often obtained from oligo-dr primed cDNA
XX CC libraries. Such ESTs are not well suited for isolating cDNA sequences
XX CC derived from the 5' ends of mRNAs and even in those cases where longer
XX CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
XX CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
XX CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
XX CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX CC They are used to obtain upstream regulatory sequences and to design
XX CC expression and secretion vectors
XX SQ Sequence 285 BP; 112 A; 49 C; 38 G; 86 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,95e+03 Length: 285
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x AAC22592 (1-285)
QY 1 GlyTyr***ValGlu 5
Db 134 GGATATACCGTGAA 120
RESULT 70
AAC22663
ID AAC22663 standard; cDNA; 286 BP.
XX AC AAC22663;
XX DT 24-JAN-1997 (first entry)
XX DE E. coli plasmid R67 dihydrofolate reductase gene.
XX KW Hypermutation; reverse transcription; biasing dNTP concentration;
XX KW Escherichia coli; dihydrofolate reductase; DHFR; plasmid R67;
XX KW polymerase chain reaction; PCR; amplification; primer; ds.
XX OS Escherichia coli.
XX FH Key Location/Qualifiers
XX FT CDS 24..260
XX FT /*tag= a
XX PN WO9617056-A1.
XX PD 06-JUN-1996.
XX PF 01-DEC-1995; 95WO-EP004749.
XX

```



```

PR 02-DEC-1994; 94EP-00402774.
XX 19-MAY-1995; 95US-00447173.
XX (INSP ) INST PASTEUR.
XX Wain-Hobson S;
XX WPI; 1996-277776/28.
XX P-PSDB; AAW02576.
XX Inducing hyper-mutation(s) in DNA or RNA - by transcribing RNA into DNA
XX in the presence of varying biased concns. of deoxy:nucleotide
XX tri:phosphate(s).
XX Example 2; Fig 5; 72pp; English.
XX Hypermutations can be introduced into RNA or DNA by reverse transcription
XX (RT) in the presence of a biasing concn. of dNTP. In an example, a
XX dihydrofolate reductase (DHFR) gene encoded by the E. coli plasmid R67
XX was utilised as the target DNA sequence. The DHFR gene was 1st amplified
XX from the plasmid using PCR and the primer pair AAT32664/65. The prod.
XX then underwent RT in a biasing concn. of dNTP, and the resulting cDNA
XX amplified using the above primer pair. The prod. was cleaved using the
XX introduced restriction sites, and ligated into the pTRC99A expression
XX vector for the transformation of E. coli cells. The cells were then
XX plated out on to standard lauria broth supplemented with trimethoprim
XX (TMP) and ampicillin. To identify the nature of the hypermutated R67
XX genes, TMP resistant colonies were grown up singly, plasmid DNA extracted
XX and the DNA sequenced and compared to the wild type DHFR DNA sequence
XX AAT32663, which encodes AAW02576
XX SQ Sequence 286 BP; 72 A; 69 C; 80 G; 65 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.96e+03 Length: 286
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x AAT32663 (1-286)

Qy 1 GlyTyr***ValGlu 5
Db 183 GGCTACGCCGTCGAG 197

RESULT 71
AAT27833
ID AAT27833 standard; cDNA; 286 BP.
XX AC AAT27833;
XX DT 24-JAN-1997 (first entry)
XX DE E. coli plasmid R67 dihydrofolate reductase gene.
XX KW Hypermutation; reverse transcription; biasing dNTP concentration;
XX KW Escherichia coli; dihydrofolate reductase; DHFR; plasmid R67;
XX KW polymerase chain reaction; PCR; amplification; primer; ds.
XX OS Escherichia coli.
XX FH Key Location/Qualifiers
XX CDS 24..260
XX FT /*tag= a
XX EP714980-A1.
XX PN 05-JUN-1996.
XX PD
XX XX 02-DEC-1994; 94EP-00402774.
XX FI

```

```

PR 02-DEC-1994; 94EP-00402774.
XX (INSP ) INST PASTEUR.
XX Martinez MA, Wain-Hobson S;
XX WPI; 1996-269713/28.
XX P-PSDB; AAR95919.
XX Introducing hyper:mutation(s) into RNA and DNA by reverse transcription -
XX in presence of biasing concn. of dNTP, also hyper:mutated nucleic acid
XX and derived mutant proteins.
XX Example 2; Fig 5; 47pp; English.
XX Hypermutations can be introduced into RNA or DNA by reverse transcription
XX (RT) in the presence of a biasing concn. of dNTP. In an example, a
XX dihydrofolate reductase (DHFR) gene encoded by the E. coli plasmid R67
XX was utilised as the target DNA sequence. The DHFR gene was 1st amplified
XX from the plasmid using PCR and the primer pair AAT27831/32. The prod.
XX then underwent RT in a biasing concn. of dNTP, and the resulting cDNA
XX amplified using the above primer pair. The prod. was cleaved using the
XX introduced restriction sites, and ligated into the pTRC99A expression
XX vector for the transformation of E. coli cells. The cells were then
XX plated out on to standard lauria broth supplemented with trimethoprim
XX (TMP) and ampicillin. To identify the nature of the hypermutated R67
XX genes, TMP resistant colonies were grown up singly, plasmid DNA extracted
XX and the DNA sequenced and compared to the wild type DHFR DNA sequence
XX AAT27833, which encodes AAR95919
XX SQ Sequence 286 BP; 72 A; 69 C; 80 G; 65 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.96e+03 Length: 286
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x AAT27833 (1-286)

Qy 1 GlyTyr***ValGlu 5
Db 183 GGCTACGCCGTCGAG 197

RESULT 72
AAH81247/c
ID AAH81247 standard; DNA; 286 BP.
XX AC AAH81247;
XX DT 21-SEP-2001 (first entry)
XX DE Escherichia coli nucleotide sequence SEQ ID NO:46.
XX KW Escherichia coli; identification; proliferation; microorganism;
XX KW antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;
XX KW bacterial growth inhibition; ds.
XX OS Escherichia coli.
XX PN WO200148209-A2.
XX PD 05-JUL-2001.
XX PF 19-DEC-2000; 2000WO-US034419.
XX PR 23-DEC-1999; 99US-0173005P.
XX PA (ELIT-) ELITRA PHARM INC.
XX FI Forsyth RA, Ohlsen KL, Zyskind JW;

```

XX WPI; 2001-457376/49.

XX Novel nucleic acids encoding proteins required for Escherichia coli

PT proliferation, useful for screening for antimicrobial agents.

XX

XX Claim 1; Page 120; 596pp; English.

XX The present invention describes a purified or isolated nucleic acid

CC sequence (I) consisting essentially of one of the 93 nucleotide sequences

CC given in AH81202 to AH81294, where expression of the nucleic acid in a

CC microorganism is capable of inhibiting proliferation of a microorganism.

CC (I) have antibacterial and antibiotic activities and can be used in gene

CC therapy. Expression of (I) in a microorganism inhibits proliferation of

CC the microorganism, and the manufactured antibiotic is useful for reducing

CC the activity or level of a gene product required for proliferation of a

CC microorganism in a subject, specifically humans. The nucleic acids that

CC inhibit bacterial growth or proliferation can be used as antisense

CC therapeutics for killing bacteria. In addition to therapeutic

CC applications, the nucleic acid sequences complementary to sequences

CC required for proliferation can be used as diagnostic tools. For example,

CC nucleic acid probes complementary to proliferation-required sequences

CC that are specific for particular species of microorganisms can be used as

CC probes to identify particular microorganism species in clinical

CC specimens. AH81295 to AH81487 encode the Escherichia coli proteins

CC given in AAG98239 to AAG98431, and AH81488 to AH81491 represent

CC oligonucleotides, which are used in the exemplification of the present

CC invention

XX

SQ Sequence 286 BP; 77 A; 72 C; 67 G; 70 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.96e+03	Length:	286
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	5	Gaps:	0

US-10-030-194A-6 (1-6) x AH81247 (1-286)

OY 1 GLYTYR\*\*\*ValGlu 5

DB 222 GGTACACCGTCGAA 208

RESULT 73

ACA13551/c

ID ACA13551 standard; DNA; 286 BP.

AC ACA13551;

XX 27-OCT-2003 (revised)

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene antisense oligonucleotide #1421.

XX Antisense; ss; prokaryotic essential gene; cell proliferation;

KW drug design.

XX Archaea.

OS WO200277183-A2.

PN 03-OCT-2002.

FD 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 1; SEQ ID NO 1421; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than S. aureus, S. typhimurium,

CC K. pneumoniae or P. aeruginosa. The present sequence is one of the 6213

CC antisense sequences of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 27-OCT-2003 to

CC standardise OS field)

XX

SQ Sequence 286 BP; 77 A; 72 C; 67 G; 70 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.96e+03	Length:	286
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	8	Gaps:	0

US-10-030-194A-6 (1-6) x ACA13551 (1-286)

OY 1 GLYTYR\*\*\*ValGlu 5

DB 222 GGTACACCGTCGAA 208

RESULT 74

AAD04460

ID AAD04460 standard; DNA; 287 BP.

XX AAD04460;

AC AAD04460;

XX 04-JUL-2001 (first entry)

DT Exon 18 of human insulin receptor gene.

XX Human; insulin receptor; exon 18; cephalic pain; therapy; headache;

KW

KW chronic paroxysmal hemicrania; vascular disorder; tension headache;  
XX migraine; single nucleotide polymorphism; SNP; ds.  
OS Homo sapiens.  
XX WO200128540-A2.  
PN 26-APR-2001.  
XX 19-OCT-2000; 2000WO-GB004051.  
PF 19-OCT-1999; 95GB-00024712.  
PR 19-OCT-1999; 99US-0160418P.  
XX (GLAX ) GLAXO GROUP LTD.  
PA Purvis IJ, McCarthy LC;  
XX WPI; 2001-290815/30.  
XX  
XX Use of agent that modulates directly or indirectly insulin receptor or  
PT insulin receptor signaling pathway in manufacture of medicament for  
PT preventing or treating cephalic pain.  
XX Disclosure; Page 35; 46pp; English.  
XX  
XX The present invention relates to the use of an agent that modulates  
CC directly or indirectly the insulin receptor or insulin receptor  
CC signalling pathway in the manufacture of a medicament for preventing or  
CC treating cephalic pain. Cephalic pain disorders are generally  
CC multifunctional disorder, with a cluster of headache, chronic paroxysmal  
CC hemicrania, headache associated with vascular disorders, headache  
CC associated with substances or their withdrawal (for example drug  
CC withdrawal), tension headache and in particular migraine with aura or  
CC migraine without aura. The treatment of cephalic pain and migraine  
CC involves manipulation of components of the glucose and lipid metabolism  
CC pathways, in particular by manipulation of the insulin receptor. Insulin  
CC receptor is an important component in the regulation of the glucose and  
CC lipid metabolism pathways. Single nucleotide polymorphisms (SNP) in the  
CC insulin receptor gene contribute to susceptibility to cephalic pain. The  
CC present sequence is exon 18 of human insulin receptor gene. SNP in this  
CC sequence contributes to susceptibility to cephalic pain  
XX  
SQ Sequence 287 BP; 51 A; 85 C; 98 G; 53 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.97e+03 Length: 287  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x AAD04460 (1-287)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 253 GGGTACTCGGTGGAG 267

RESULT 75  
AAH31260  
ID AAH31260 standard; DNA; 287 BP.  
XX  
XX AAH31260;  
XX  
XX 25-JUL-2001 (first entry)  
XX  
XX Human insulin receptor gene exon 18.  
XX  
XX Human; insulin receptor; antimigraine; cephalic pain; diagnosis;  
KW susceptibility; single nucleotide polymorphism; SNP; SNP detection;  
XX migraine; ds.  
XX

OS Homo sapiens.  
XX WO200129256-A2.  
XX  
XX 26-APR-2001.  
XX 19-OCT-2000; 2000WO-GB004050.  
PF 19-OCT-1999; 99GB-00024717.  
PR (GLAX ) GLAXO GROUP LTD.  
PA Purvis IJ, McCarthy LC;  
XX WPI; 2001-316247/33.  
XX  
XX Diagnosing susceptibility to cephalic pain such as migraine by typing  
PT insulin receptor gene or protein in vivo, or in a biological sample and  
PT determining individual's susceptibility to cephalic pain.  
XX  
XX Disclosure; Page 36; 46pp; English.  
XX  
XX The present sequence is provided in a specification relating to a method  
CC for diagnosing susceptibility to cephalic pain in an individual. The  
CC method comprises typing in vivo or in a sample from the individual, the  
CC insulin receptor gene region or insulin receptor protein and thus  
CC determining whether the individual is susceptible to cephalic pain.  
CC Polymorphisms in the insulin receptor gene that cause susceptibility to  
CC cephalic pain have been found. Susceptibility may be diagnosed using a  
CC probe, primer or antibody which is capable of detecting an insulin  
CC receptor gene region or insulin receptor protein polymorphism. The method  
CC is useful for diagnosing susceptibility to migraine, cluster headache,  
CC chronic paroxysmal hemicrania, headache associated with vascular  
CC disorders, headache associated with substances or their withdrawal,  
CC tension headache, and so on. It is useful for assessing the efficacy of  
CC agents in relieving cephalic pain, and can be used to assess the ability  
CC of agents to modulate insulin receptor signalling activity. The method  
CC may also be used to assess the predisposition and/or susceptibility of an  
CC individual to the development of diseases mediated by the insulin  
CC receptor and in the development of new drug therapies which selectively  
CC target one or more allelic variants of the insulin receptor gene  
XX  
SQ Sequence 287 BP; 51 A; 85 C; 98 G; 53 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.97e+03 Length: 287  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x AAH31260 (1-287)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 253 GGGTACTCGGTGGAG 267

Search completed: November 3, 2004, 15:20:50  
Job time : 351 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 3, 2004, 13:32:27 ; Search time 2509 Seconds  
(without alignments)  
87.142 Million cell updates/sec

Title: US-10-030-194A-6

Perfect score: 24

Sequence: 1 GYXVEX 6

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO gpool /P/US10030194/runat\_01112004.184847.28965/app\_query.fasta\_1.398  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10030194 @CGN 1 1 7873 @runat\_01112004.184847.28965 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_hc:\*
- 4: gb\_est3:\*
- 5: gb\_est4:\*
- 6: gb\_est5:\*
- 7: gb\_est6:\*
- 8: gb\_gsl:\*
- 9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	91.7	58	AA232607	AA232607 zr28d08.r
2	22	91.7	72	AZ389716	AZ389716 1M0150H12
3	22	91.7	81	AL763170	AL763170 Arabidops
4	22	91.7	83	CG584681	CG584681 OST230973
5	22	91.7	92	AZ500158	AZ500158 1M0338C23
6	22	91.7	92	AZ773696	AZ773696 2M0001N15
7	22	91.7	98	CF099395	CF099395 rd72d05.y
8	22	91.7	101	TA367A02Q	TA367A02Q T. brucei
9	22	91.7	101	AG203579	AG203579 Pan trogl

C	10	22	91.7	105	1	AA249463	AA249463 j5324.seq
C	11	22	91.7	106	1	AA393370	AA393370 tct57c11.s
C	12	22	91.7	106	4	BI864792	BI864792 tct96n07.y
C	13	22	91.7	106	7	CK256244	CK256244 EST739881
C	14	22	91.7	106	7	CNO27671	CNO27671 UMC-p4c1v
C	15	22	91.7	108	9	CC649660	CC649660 OGLBD35TH
C	16	22	91.7	110	8	AZ309801	AZ309801 1M0017M04
C	17	22	91.7	114	1	AI664163	AI664163 ue79g10.x
C	18	22	91.7	116	6	CA353391	CA353391 624896 NC
C	19	22	91.7	118	1	AI906955	AI906955 QV-BT130-
C	20	22	91.7	118	5	BUI16086	BUI16086 N060C10 P
C	21	22	91.7	126	2	AW935423	AW935423 CM3-DT000
C	22	22	91.7	127	8	AZ083244	AZ083244 RPCI-23-4
C	23	22	91.7	128	1	AJ770527	AJ770527 AJ770527
C	24	22	91.7	129	4	BG994255	BG994255 CM3-HT115
C	25	22	91.7	129	8	CC356105	CC356105 PUHPV81TD
C	26	22	91.7	129	9	AL766514	AL766514 Arabidops
C	27	22	91.7	129	9	CG644531	CG644531 OST386390
C	28	22	91.7	130	4	BG140779	BG140779 EST481221
C	29	22	91.7	130	4	BG385119	BG385119 306732 MA
C	30	22	91.7	131	2	AW750597	AW750597 PM3-CN002
C	31	22	91.7	131	5	BQ976013	BQ976013 QH120D04.
C	32	22	91.7	131	5	BQ978589	BQ978589 QH15H23.Y
C	33	22	91.7	131	6	CD297327	CD297327 StrPu691.
C	34	22	91.7	131	6	CF088605	CF088605 QHM1F08.Y
C	35	22	91.7	131	8	AZ752194	AZ752194 RPCI-24-1
C	36	22	91.7	132	2	BF557111	BF557111 UI-R-CO-9
C	37	22	91.7	132	4	BI843879	BI843879 ft91e03.Y
C	38	22	91.7	132	5	BU020428	BU020428 QHE27G13.
C	39	22	91.7	132	9	CG653419	CG653419 OST418752
C	40	22	91.7	134	1	AU009303	AU009303 AU009303
C	41	22	91.7	134	1	AU009304	AU009304 AU009304
C	42	22	91.7	136	5	BQ293027	BQ293027 MRI-AN003
C	43	22	91.7	138	1	AI053196	AI053196 TENU1620
C	44	22	91.7	138	5	BQ912701	BQ912701 QHA22N01.
C	45	22	91.7	138	5	BW020881	BW020881 BW020881
C	46	22	91.7	140	2	AW750578	AW750578 RC0-CN002
C	47	22	91.7	141	2	AW582529	AW582529 RC2-ST025
C	48	22	91.7	141	4	BG930246	BG930246 f231-148
C	49	22	91.7	141	6	CF023361	CF023361 QBQ3e12.x
C	50	22	91.7	142	2	AW380229	AW380229 QV3-HT026
C	51	22	91.7	142	9	CG602574	CG602574 OST276104
C	52	22	91.7	143	2	BE161529	BE161529 RC3-HT044
C	53	22	91.7	143	7	X83334	X83334 MMR542UNI M
C	54	22	91.7	144	2	BF843378	BF843378 CM4-HT115
C	55	22	91.7	144	2	AW300193	AW300193 xs60e05.x
C	56	22	91.7	144	2	AW577008	AW577008 PM3-BT034
C	57	22	91.7	144	8	BH438734	BH438734 BOHOS14TF
C	58	22	91.7	146	2	BF368176	BF368176 CM3-GN004
C	59	22	91.7	148	1	AV771220	AV771220 AV771220
C	60	22	91.7	148	2	BE654176	BE654176 UI-M-ANI-
C	61	22	91.7	148	8	AZ461820	AZ461820 1M0267A24
C	62	22	91.7	148	8	AZ749976	AZ749976 RPCI-24-7
C	63	22	91.7	150	6	CB817900	CB817900 a6d02p2.x
C	64	22	91.7	151	1	AI459749	AI459749 ap16h12.x
C	65	22	91.7	151	2	BF047816	BF047816 dc8se04.y
C	66	22	91.7	152	1	AV017443	AV017443 AV017443
C	67	22	91.7	152	7	N76484	N76484 Yz87h10.r1
C	68	22	91.7	153	4	BI045919	BI045919 MR3-FN020
C	69	22	91.7	153	5	BQ255119	BQ255119 NISC_jp12
C	70	22	91.7	153	9	CR066309	CR066309 Reverse.s
C	71	22	91.7	155	1	AV950897	AV950897 AV950897
C	72	22	91.7	155	2	BF806750	BF806750 RC2-C1008
C	73	22	91.7	155	2	BE154383	BE154383 PM2-HT034
C	74	22	91.7	156	5	BM986915	BM986915 UI-H-COO-
C	75	22	91.7	156	7	CN945347	CN945347 011010AVB
C	76	22	91.7	156	9	CE309035	CE309035 tigr-gss-
C	77	22	91.7	157	7	CO264091	CO264091 EKJ37354.
C	78	22	91.7	157	7	W83329	W83329 mf27h06.r1
C	79	22	91.7	158	1	AI035926	AI035926 uba9e06.r
C	80	22	91.7	158	1	AA624439	AA624439 vn78c11.r
C	81	22	91.7	158	2	AW947602	AW947602 RC0-MT000
C	82	22	91.7	159	2	AW486029	AW486029 RC9371 MAR



229	22	91.7	217	6	CB220464	CR220464	1AB022B07	BB183020	BB183020
230	22	91.7	218	2	AW224204	AW224204	EST301111	B0643512	B0643512
231	22	91.7	218	2	BE155658	BE155658	FM2-HT035	CD270767	CD270767
232	22	91.7	218	5	BU023155	BU023155	QHE9L24.Y	F05093	F05093
233	22	91.7	218	5	CU524869	CU524869	AL0036.Sa	BY781132	BY781132
234	22	91.7	219	9	CB699946	CB699946	tigr-g8s-	CB279993	CB279993
235	22	91.7	220	1	AJ502951	AJ502951	AJ502951	CR516307	CR516307
236	22	91.7	220	1	BF509543	BF509543	UI-H-B14-	BH663496	BH663496
237	22	91.7	221	7	CK976704	CK976704	4107778.B	AG268433	AG268433
238	22	91.7	221	8	AZ282067	AZ282067	RPCI-23-1	BB010831	BB010831
239	22	91.7	222	8	BB608574	BB608574	sq38a06.Y	AU059033	AU059033
240	22	91.7	222	9	CB770910	CB770910	tigr-g8s-	AW416558	AW416558
241	22	91.7	223	1	AL828457	AL828457	AL828457	AM483164	AM483164
242	22	91.7	223	1	AV351438	AV351438	AV351438	BJ226986	BJ226986
243	22	91.7	223	9	CE293953	CE293953	tigr-g8s-	BZ826819	BZ826819
244	22	91.7	224	2	BB184677	BB184677	BB184677	AA813661	AA813661
245	22	91.7	225	2	BZ484957	BZ484957	BZ484957	AA119863	AA119863
246	22	91.7	225	6	CB279563	CB279563	ru83e04.Y	AA301321	AA301321
247	22	91.7	226	1	AV588617	AV588617	AV588617	BB137060	BB137060
248	22	91.7	227	2	BB539497	BB539497	BB539497	BG730825	BG730825
249	22	91.7	227	8	AZ211564	AZ211564	SP_0155.B	BG991998	BG991998
250	22	91.7	228	2	BB063514	BB063514	BB063514	BM411299	BM411299
251	22	91.7	228	2	BB379542	BB379542	BB379542	AV426524	AV426524
252	22	91.7	228	5	BU020431	BU020431	QHE27G17	BB147894	BB147894
253	22	91.7	228	8	BZ887084	BZ887084	CH240_219	BZ887084	BZ887084
254	22	91.7	229	2	BF798533	BF798533	CM0-CI006	BP735760	BP735760
255	22	91.7	229	2	BE113704	BE113704	UI-R-BJ1-	BN861993	BN861993
256	22	91.7	229	5	BX918147	BX918147	BX918147	BF748119	BF748119
257	22	91.7	230	2	BB254444	BB254444	BB254444	BF758163	BF758163
258	22	91.7	230	4	BT043940	BT043940	MR4-OT016	BF821814	BF821814
259	22	91.7	230	5	BQ311533	BQ311533	RC6-BN027	BF821814	BF821814
260	22	91.7	231	2	BF801591	BF801591	CM1-CI009	CE425463	CE425463
261	22	91.7	231	7	CK524061	CK524061	rswea0_01	AV367099	AV367099
262	22	91.7	231	7	CO487448	CO487448	GQ0227_B7	AA3379073	AA3379073
263	22	91.7	231	7	W06798	W06798	SMEST0434.S	BF909261	BF909261
264	22	91.7	231	8	AZ484420	AZ484420	1M0310B22	AM477834	AM477834
265	22	91.7	231	8	BZ090362	BZ090362	CH230-214	CK437988	CK437988
266	22	91.7	232	1	A1065843	A1065843	ag94f03.x	CM106641	CM106641
267	22	91.7	232	2	AW221616	AW221616	EST298427	CN929641	CN929641
268	22	91.7	232	5	BU018501	BU018501	QHE18N13	CO602142	CO602142
269	22	91.7	232	6	CD962998	CD962998	SDS_129.G	AG229502	AG229502
270	22	91.7	232	7	F07889	F07889	HSC2KG051.n	AA932973	AA932973
271	22	91.7	233	1	AV352303	AV352303	AV352303	BP103793	BP103793
272	22	91.7	233	2	BB059640	BB059640	BB059640	CE677884	CE677884
273	22	91.7	233	9	CG670511	CG670511	OGJBR21TV	AV208382	AV208382
274	22	91.7	233	9	CE242325	CE242325	tigr-g8s-	AA494608	AA494608
275	22	91.7	235	2	BB268666	BB268666	BB268666	BE077269	BE077269
276	22	91.7	235	5	BU227723	BU227723	603399589	CF010085	CF010085
277	22	91.7	236	2	BF380983	BF380983	CM4-UT007	AV283631	AV283631
278	22	91.7	236	2	BB035248	BB035248	BB035248	CA051179	CA051179
279	22	91.7	236	7	CO363766	CO363766	RTK1_11.E	CA796270	CA796270
280	22	91.7	236	7	CO537220	CO537220	tab67h08.	BF947046	BF947046
281	22	91.7	236	9	CG500168	CG500168	OST42918	BB854154	BB854154
282	22	91.7	237	1	AV010521	AV010521	AV010521	CR522376	CR522376
283	22	91.7	237	2	BB171903	BB171903	BB171903	BB575724	BB575724
284	22	91.7	237	2	BB843515	BB843515	CM3-TN006	BM207927	BM207927
285	22	91.7	237	4	BT045943	BT045943	MR3-FN020	CF242715	CF242715
286	22	91.7	237	9	CG608803	CG608803	OST289311	BX989968	BX989968
287	22	91.7	237	9	CG661787	CG661787	OST443443	EX998597	EX998597
288	22	91.7	238	1	AV029409	AV029409	AV029409	CC848224	CC848224
289	22	91.7	238	1	BB003162	BB003162	BB003162	BI254921	BI254921
290	22	91.7	238	2	BB019656	BB019656	BB019656	CF093100	CF093100
291	22	91.7	238	2	BB057298	BB057298	BB057298	BB215749	BB215749
292	22	91.7	238	5	BW199817	BW199817	BW199817	BB368829	BB368829
293	22	91.7	238	7	CN115656	CN115656	EC3CAA43A	BW265140	BW265140
294	22	91.7	238	9	CB656762	CB656762	tigr-g8s-	CA568277	CA568277
295	22	91.7	238	9	CG648410	CG648410	OST401727	CF089540	CF089540
296	22	91.7	239	2	BF095791	BF095791	CM4-UT007	AA097816	AA097816
297	22	91.7	239	5	BQ968398	BQ968398	QHE33M10	AA324104	AA324104
298	22	91.7	239	7	CK338773	CK338773	CO808D06-	BB152647	BB152647
299	22	91.7	239	7	CK763993	CK763993	ltu01-12m	AQ107496	AQ107496
300	22	91.7	240	1	AU072247	AU072247	hug55a03.x	AA966572	AA966572
301	22	91.7	240	2	AW594091	AW594091	AW594091	BF849864	BF849864

375	22	91.7	257	1	AU272824	AU272824	AU272824	C 448	22	91.7	271	1	AV278603	AV278603
376	22	91.7	258	2	BF593846	BF593846	BF593846	449	22	91.7	271	2	AW193052	AW193052
377	22	91.7	259	3	BF605310	BF605310	BF605310	C 450	22	91.7	271	2	AW196457	AW196457
378	22	91.7	258	4	BJ047769	BJ047769	BJ047769	451	22	91.7	271	2	AW578969	AW578969
379	22	91.7	258	7	F22978	F22978	F22978	452	22	91.7	271	2	AW938690	AW938690
380	22	91.7	258	1	AU304776	AU304776	AU304776	C 453	22	91.7	271	2	BB529585	BB529585
381	22	91.7	259	4	BF987643	BF987643	BF987643	454	22	91.7	271	2	BE717881	BE717881
382	22	91.7	259	9	CC936165	CC936165	CC936165	455	22	91.7	271	6	CB100499	CB100499
383	22	91.7	260	2	BB595329	BB595329	BB595329	C 456	22	91.7	271	8	AZ931567	AZ931567
384	22	91.7	260	8	AF114702	AF114702	AF114702	C 457	22	91.7	271	2	AV074609	AV074609
385	22	91.7	260	8	AZ224469	AZ224469	AZ224469	C 458	22	91.7	272	2	BB416240	BB416240
386	22	91.7	261	4	BI078897	BI078897	BI078897	C 459	22	91.7	272	2	BE717883	BE717883
387	22	91.7	261	4	BI078790	BI078790	BI078790	461	22	91.7	272	8	AZ885801	AZ885801
388	22	91.7	261	5	CUK996147	CUK996147	CUK996147	C 462	22	91.7	272	8	CG933682	CG933682
389	22	91.7	261	5	BU984914	BU984914	BU984914	C 463	22	91.7	273	2	BB456672	BB456672
390	22	91.7	261	8	BZ149303	BZ149303	BZ149303	C 464	22	91.7	273	2	BB597143	BB597143
391	22	91.7	261	8	BZ160573	BZ160573	BZ160573	465	22	91.7	273	2	BB598742	BB598742
392	22	91.7	261	9	CG505068	CG505068	CG505068	466	22	91.7	273	2	BB597143	BB597143
393	22	91.7	262	1	AJ301152	AJ301152	AJ301152	467	22	91.7	273	9	CL442574	CL442574
394	22	91.7	262	5	BQ032701	BQ032701	BQ032701	C 468	22	91.7	274	2	BF530231	BF530231
395	22	91.7	262	9	FR003176	FR003176	FR003176	C 469	22	91.7	274	2	BF530231	BF530231
396	22	91.7	262	9	CR521180	CR521180	CR521180	C 470	22	91.7	274	2	BF530231	BF530231
397	22	91.7	262	9	FR003176	FR003176	FR003176	C 471	22	91.7	274	2	BF530231	BF530231
398	22	91.7	263	1	AI986052	AI986052	AI986052	C 472	22	91.7	274	2	BF530231	BF530231
399	22	91.7	263	1	AV283284	AV283284	AV283284	C 473	22	91.7	274	2	BF530231	BF530231
400	22	91.7	263	4	BM151259	BM151259	BM151259	C 474	22	91.7	274	2	BF530231	BF530231
401	22	91.7	263	4	CF777390	CF777390	CF777390	C 475	22	91.7	274	2	BF530231	BF530231
402	22	91.7	263	9	CL223428	CL223428	CL223428	C 476	22	91.7	274	2	BF530231	BF530231
403	22	91.7	263	9	CL223428	CL223428	CL223428	C 477	22	91.7	274	2	BF530231	BF530231
404	22													



C 521	22	91.7	280	4	B1318300	B1318300 fq70f02.y	594	22	91.7	291	7	CF531305	UI-M-FY0-
C 522	22	91.7	280	6	C0087142	C0087142 MC1-0035T	595	22	91.7	292	1	AA862422	AA862422 z95f02.s
C 523	22	91.7	280	8	A2800415	A2800415 2M0058H21	596	22	91.7	292	1	AA452427	AA452427 xx5f11.f
C 524	22	91.7	280	8	CC183486	CC183486 XE426 Bay	597	22	91.7	292	1	AA508693	AA508693 n122d12.s
C 525	22	91.7	281	2	B8087297	B8087297 BB087297	598	22	91.7	292	2	BB070544	BB070544 BB070544
C 526	22	91.7	281	2	B8240892	B8240892 BB240892	C 599	22	91.7	292	2	BF091660	BF091660 MRI-TN004
C 527	22	91.7	281	2	B8483136	B8483136 BB483136	C 600	22	91.7	292	2	BP742043	BP742043 BP742043
C 528	22	91.7	281	2	B8542480	B8542480 BB542480	601	22	91.7	293	1	AA862473	AA862473 Q996f02.s
C 529	22	91.7	281	5	BP702840	BP702840 BP702840	C 602	22	91.7	293	2	BB513953	BB513953 BB513953
C 530	22	91.7	281	5	BQ911194	BQ911194 QHA16G21.	C 603	22	91.7	293	2	BB719158	BB719158 BB719158
C 531	22	91.7	281	7	M478789	M478789 EST00937.Hi	604	22	91.7	293	4	BI046607	BI046607 MR3-FN020
C 532	22	91.7	281	7	CF394192	CF394192 RTDS2.4.C	C 605	22	91.7	293	4	BJ354384	BJ354384 BJ354384
C 533	22	91.7	281	8	A2375091	A2375091 1M0128J16	C 606	22	91.7	293	5	BM276294	BM276294 BM276294
C 534	22	91.7	281	9	CR198457	CR198457 Forward s	607	22	91.7	293	8	CF403527	CF403527 CSECS010C
C 535	22	91.7	282	1	AV120357	AV120357 AV120357	608	22	91.7	293	8	AZ068747	AZ068747 RPCI-23-4
C 536	22	91.7	282	7	Z24880	Z24880 HSB66A112.S	609	22	91.7	293	8	BZ627751	BZ627751 ih54g07.g
C 537	22	91.7	282	9	CL230194	CL230194 ZWMBBC007	610	22	91.7	294	2	AW037149	AW037149 614018F02
C 538	22	91.7	283	1	A1887133	A1887133 w148e12.x	611	22	91.7	294	7	N79048	N79048 zb48a10.g1
C 539	22	91.7	283	1	AA360663	AA360663 EST69892	612	22	91.7	294	7	RJ1875	RJ1875 yH59f09.x1
C 540	22	91.7	283	2	BF380986	BF380986 CM4-UT007	613	22	91.7	295	1	AV423395	AV423395 AV423395
C 541	22	91.7	283	2	BB367053	BB367053 BB367053	614	22	91.7	295	2	BB142499	BB142499 BB142499
C 542	22	91.7	283	2	BE000320	BE000320 MR0-BN007	615	22	91.7	295	7	CK248536	CK248536 EST732173
C 543	22	91.7	283	2	BE174130	BE174130 QV1-HT057	616	22	91.7	295	7	CN197541	CN197541 TgESTzy14
C 544	22	91.7	283	7	CF394192	CF394192 RTDS2.4.C	617	22	91.7	295	8	AQ096602	AQ096602 HS_3038.A
C 545	22	91.7	284	1	AA090892	AA090892 Y07977.se	C 618	22	91.7	296	1	AJ494119	AJ494119 AJ394113
C 546	22	91.7	284	1	AV131396	AV131396 AV131396	C 619	22	91.7	296	2	BB551286	BB551286 BB551286
C 547	22	91.7	284	1	AV343193	AV343193 AV343193	620	22	91.7	296	5	BQ325887	BQ325887 RC3-CI008
C 548	22	91.7	284	2	BB538662	BB538662 BB538662	C 621	22	91.7	296	7	CO687399	CO687399 DG11-230P
C 549	22	91.7	284	2	BB5690848	BB5690848 uu62d03.x	C 622	22	91.7	297	2	BB229782	BB229782 BB229782
C 550	22	91.7	284	5	BN146658	BN146658 BN146658	C 623	22	91.7	297	2	BB339215	BB339215 BB339215
C 551	22	91.7	284	6	CA911845	CA911845 PCSC08232	C 624	22	91.7	297	6	CB701613	CB701613 AMGNNUC.U
C 552	22	91.7	284	6	CB101427	CB101427 k171d10.y	625	22	91.7	297	6	CD729934	CD729934 4038047.1
C 553	22	91.7	284	9	AC042433	AC042433 RPCI-23-2	626	22	91.7	298	1	BZ617727	BZ617727 ig21d04.b
C 554	22	91.7	284	9	CG602470	CG602470 ZWMBBC041	C 627	22	91.7	298	1	AI644063	AI644063 vu85c06.x
C 555	22	91.7	284	9	CG502243	CG502243 OST46996	C 628	22	91.7	298	2	BF885450	BF885450 MR3-TN016
C 556	22	91.7	285	1	AV035881	AV035881 AV035881	C 629	22	91.7	298	2	AW086727	AW086727 g04e02.x
C 557	22	91.7	285	1	BB193616	BB193616 BB193616	630	22	91.7	298	2	AW789651	AW789651 Co4256-F
C 558	22	91.7	285	2	BB292893	BB292893 BB292893	631	22	91.7	298	2	BB111692	BB111692 BB111692
C 559	22	91.7	285	6	CB870783	CB870783 HC15E12w	632	22	91.7	298	2	BB292697	BB292697 BB292697
C 560	22	91.7	285	8	A2931626	A2931626 474.dhz86	633	22	91.7	298	2	BB340490	BB340490 BB340490
C 561	22	91.7	285	8	BH031800	BH031800 RPCI-24-2	C 634	22	91.7	298	2	BB718935	BB718935 BB718935
C 562	22	91.7	286	1	AA911697	AA911697 ok88e09.s	635	22	91.7	298	6	CA516999	CA516999 KS09068E1
C 563	22	91.7	286	2	AW352208	AW352208 CM3-HT013	636	22	91.7	298	7	CO055199	CO055199 TgESTzy02
C 564	22	91.7	286	2	BB369853	BB369853 BB369853	C 637	22	91.7	299	1	AA077651	AA077651 7B35C09.C
C 565	22	91.7	286	2	BB576864	BB576864 L0-1671T3	638	22	91.7	299	2	BF521196	BF521196 EST458743
C 566	22	91.7	286	8	AO445813	AO445813 nbxb0056F	639	22	91.7	299	2	BF521281	BF521281 EST458692
C 567	22	91.7	287	5	BW019672	BW019672 QHE22M17.	640	22	91.7	299	2	AW312789	AW312789 5093.MARC
C 568	22	91.7	287	6	CD295206	CD295206 StzPu691.	C 641	22	91.7	299	2	BB060254	BB060254 BB060254
C 569	22	91.7	287	6	CD931728	CD931728 GR45.115G	642	22	91.7	299	2	BB136052	BB136052 BB136052
C 570	22	91.7	287	6	CF034190	CF034190 QCF2c02.y	C 643	22	91.7	299	5	BP582018	BP582018 BP582018
C 571	22	91.7	287	7	CM197542	CM197542 TgESTzy14	C 644	22	91.7	299	7	T13713	T13713 1878.Lambda
C 572	22	91.7	287	9	CB821722	CB821722 tigr-ges-	C 645	22	91.7	300	1	AA051900	AA051900 zEST00592
C 573	22	91.7	287	9	CG502254	CG502254 OST47010	646	22	91.7	300	1	AJ477449	AJ477449 AJ477449
C 574	22	91.7	288	1	AA343544	AA343544 EST49339	647	22	91.7	300	1	AJ479003	AJ479003 AJ479003
C 575	22	91.7	288	2	BF393049	BF393049 UI-R-CA0-	648	22	91.7	300	1	AJ479013	AJ479013 AJ479013
C 576	22	91.7	288	2	BB355198	BB355198 BB355198	649	22	91.7	300	1	AA235018	AA235018 z638c11.s
C 577	22	91.7	288	6	CL6351	CL6351 C16351 Clon	650	22	91.7	300	1	BB142543	BB142543 BB142543
C 578	22	91.7	288	7	N72377	N72377 yv38e11.r1	651	22	91.7	300	5	BU017825	BU017825 QHE16P11.
C 579	22	91.7	289	1	AI423637	AI423637 tf61e03.x	C 652	22	91.7	300	8	AZ755516	AZ755516 CQ12h06.f
C 580	22	91.7	289	1	AA163883	AA163883 mr19h03.f	C 653	22	91.7	300	8	AZ833427	AZ833427 2M0115D06
C 581	22	91.7	289	1	AV144507	AV144507 AV144507	C 654	22	91.7	301	1	AV337830	AV337830 AV337830
C 582	22	91.7	289	2	BB033165	BB033165 BB033165	655	22	91.7	301	4	BG138531	BG138531 EST478973
C 583	22	91.7	289	2	BB538663	BB538663 BB538663	C 656	22	91.7	301	5	BP648020	BP648020 BP648020
C 584	22	91.7	289	2	BB599819	BB599819 BB599819	657	22	91.7	301	5	BU122497	BU122497 D03146Q33
C 585	22	91.7	289	2	BB599819	BB599819 UI-M-BZ1-	C 658	22	91.7	301	8	BU964252	BU964252 EST384.He
C 586	22	91.7	289	8	AZ891712	AZ891712 RPCI-24-1	659	22	91.7	301	1	AV337830	AV337830 AV337830
C 587	22	91.7	289	8	CC106627	CC106627 CSU-K34.1	660	22	91.7	301	5	BP648020	BP648020 BP648020
C 588	22	91.7	290	1	AA483899	AA483899 ne92a06.s	C 661	22	91.7	301	5	BU122497	BU122497 D03146Q33
C 589	22	91.7	290	2	BF909264	BF909264 CM4-UT007	662	22	91.7	301	7	CO634808	CO634808 DG9-95h1
C 590	22	91.7	290	2	BB415145	BB415145 BB415145	C 663	22	91.7	301	8	BZ613437	BZ613437 KBCH139M1
C 591	22	91.7	290	6	CB102012	CB102012 k176e03.y	C 664	22	91.7	301	8	BZ613437	BZ613437 KBCH139M1
C 592	22	91.7	291	1	AI159522	AI159522 z80b12.f	C 665	22	91.7	302	1	AA735202	AA735202 LD01579.5
C 593	22	91.7	291	2	BB043295	BB043295 BB043295	C 666	22	91.7	302	2	BB254640	BB254640 BB254640



813	22	91.7	322	2	BE327051	hw06d06.x	886	22	91.7	333	4	BJ735507	BJ735507
814	22	91.7	322	2	B1807658	A007E01.0	887	22	91.7	333	5	B0697870	B0697870
815	22	91.7	322	6	CD736806	4020688.1	c 888	22	91.7	333	6	CB909937	USDA-FB.1
c 816	22	91.7	322	9	CR151725	Forward.s	c 889	22	91.7	333	6	CD476039	na003-2ms
817	22	91.7	323	1	AA905702	ok02h02.s	c 890	22	91.7	333	7	CK529486	rswfao.01
818	22	91.7	323	1	AJ539819	AJ539819	c 891	22	91.7	333	7	CK529486	rswfao.01
819	22	91.7	323	2	BE503130	BE503130	c 892	22	91.7	333	7	CO601579	DG8-223a2
c 820	22	91.7	323	2	B8697449	CM1-CT042	c 893	22	91.7	333	7	R31872	vh59e10.x1
c 821	22	91.7	323	4	BG008207	CM2-GN022	c 894	22	91.7	333	8	AQ993315	RPCI-23-3
c 822	22	91.7	323	6	BY778578	BY778578	c 895	22	91.7	333	9	CR101246	Reverse.s
c 823	22	91.7	323	6	CD588479	RK047ALE0	c 896	22	91.7	334	2	BF322121	maa34d04.
c 824	22	91.7	323	8	A2475205	1M0293H13	c 897	22	91.7	335	2	BF704102	MI-P-O1-a
c 825	22	91.7	323	9	AG230126	Lotus.cor	c 898	22	91.7	335	4	BG589599	EST497441
c 826	22	91.7	324	2	B8208204	BB208204	c 899	22	91.7	335	5	BQ261631	fz67c04.y
c 827	22	91.7	324	2	BS051782	BS051782	c 900	22	91.7	335	5	BQ284237	fz83e01.y
c 828	22	91.7	324	5	BQ910364	QHA13P08.	c 901	22	91.7	335	5	BQ974743	QHL16134.
c 829	22	91.7	324	7	BP692095	BP692095	c 902	22	91.7	335	7	CF578321	AGENCOURT
c 830	22	91.7	325	1	AV825566	AV825566	c 903	22	91.7	335	8	AZ234394	RPCI-23-6
c 831	22	91.7	325	1	BB140173	BB140173	c 904	22	91.7	335	9	CE056584	tigr-g88-
c 832	22	91.7	325	5	BY305828	BY305828	c 905	22	91.7	336	1	AU223776	tigr-g88-
c 833	22	91.7	325	5	BY517741	BY517741	c 906	22	91.7	336	1	AA287011	z857b04.s
c 834	22	91.7	325	6	CB967900	egx25c06	c 907	22	91.7	336	4	BG022990	daa22f05.
c 835	22	91.7	325	9	A2247167	RPCI-23-3	c 908	22	91.7	336	4	BG556507	EMI.38.A0
c 836	22	91.7	325	9	CE634194	tigr-g88-	c 909	22	91.7	336	4	BI401576	MI-P-CFO-
c 837	22	91.7	326	4	BG038747	dg24c07.x	c 910	22	91.7	336	6	CD089338	CD089338
c 838	22	91.7	326	7	CN931353	000427AFB	c 911	22	91.7	336	7	CF499911	MS1-0059T
c 839	22	91.7	327	2	AW297321	UI-H-BWO-	c 912	22	91.7	336	8	AZ930554	474.dh255
c 840	22	91.7	327	6	BA387864	BA387864	c 913	22	91.7	337	1	AV047614	AV047614
c 841	22	91.7	327	6	CA785826	sat39g07.	c 914	22	91.7	337	2	AW191855	x176c08.x
c 842	22	91.7	327	7	CN760847	ID0AA29C	c 915	22	91.7	337	2	BB552525	BB552525
c 843	22	91.7	327	8	AQ090068	HS.3006.B	c 916	22	91.7	337	4	BG486051	da810g08.
c 844	22	91.7	328	1	AA896889	vy12b09.r	c 917	22	91.7	337	4	EM688825	UI-E-CQ1-
c 845	22	91.7	328	1	A1910502	wf82g11.x	c 918	22	91.7	337	5	BU016795	QHE14A15.
c 846	22	91.7	328	5	BQ835321	kk70f07.y	c 919	22	91.7	337	5	BY103035	BY103035
c 847	22	91.7	328	7	CN612916	CsgEST00	c 920	22	91.7	337	8	BZ156754	CH230-312
c 848	22	91.7	328	7	CN849814	000917AAF	c 921	22	91.7	337	9	CL608043	CH240.174
c 849	22	91.7	328	9	CE399375	tigr-g88-	c 922	22	91.7	338	1	AI141691	ot08e01.x
c 850	22	91.7	329	1	AA984342	am84d01.s	c 923	22	91.7	338	1	AV014984	AV014984
c 851	22	91.7	329	1	A1969180	wx36f11.x	c 924	22	91.7	338	1	BF989730	IL5-GN017
c 852	22	91.7	329	4	BI070988	C050P47U	c 925	22	91.7	338	4	BF989730	IL5-GN017
c 853	22	91.7	329	6	CA911844	PCSC07759	c 926	22	91.7	339	4	BJ542135	BJ542135
c 854	22	91.7	329	6	CB100150	kl50a10.y	c 927	22	91.7	339	5	BU017766	QHE16M20.
c 855	22	91.7	329	7	CK091915	G57P61.3	c 928	22	91.7	339	6	CB694666	AMGNUG.N
c 856	22	91.7	329	7	CK674513	ZF101-P00	c 929	22	91.7	339	6	AL885953	AL885953
c 857	22	91.7	329	8	CO610405	DG8-93k14	c 930	22	91.7	339	6	BF989729	IL39-GN017
c 858	22	91.7	329	8	A2444144	1M0239A24	c 931	22	91.7	339	6	BJ542135	BJ542135
c 859	22	91.7	330	1	AA023248	mh69h08.r	c 932	22	91.7	339	6	BU017766	QHE16M20.
c 860	22	91.7	330	1	AL010846	EST205297	c 933	22	91.7	339	6	CB694666	AMGNUG.N
c 861	22	91.7	330	1	AL787156	AL787156	c 934	22	91.7	339	6	CD405609	Gm.ck2331
c 862	22	91.7	330	1	AV976812	AV976812	c 935	22	91.7	339	6	CK343267	K0549A12-
c 863	22	91.7	330	1	AW425637	57869.MAR	c 936	22	91.7	339	6	AI362562	qv17a09.x
c 864	22	91.7	330	6	CA734821	wp11s.pk0	c 937	22	91.7	340	1	AI362562	qv17a09.x
c 865	22	91.7	330	6	CD118680	ME1-0050T	c 938	22	91.7	340	2	AW904900	RC5-NN106
c 866	22	91.7	330	8	AZ868675	2M0180B04	c 939	22	91.7	340	2	BE631856	uu55e09.y
c 867	22	91.7	331	1	AL871566	we02g11.x	c 940	22	91.7	340	2	CA795539	Cac.BL.25
c 868	22	91.7	331	1	AV018063	AV018063	c 941	22	91.7	340	2	CF150635	AGENCOURT
c 869	22	91.7	331	2	BF518115	NXSI.035	c 942	22	91.7	340	2	CK340067	CO880G02-
c 870	22	91.7	331	2	BB095482	BB095482-	c 943	22	91.7	340	2	CK340067	CO880G02-
c 871	22	91.7	331	4	BB484634	231938.BA	c 944	22	91.7	340	2	AZ638934	1M0499B07
c 872	22	91.7	331	4	BJ489664	BJ489664	c 945	22	91.7	340	2	CE614248	tigr-g88-
c 873	22	91.7	331	5	B0643511	mgmk016xK	c 946	22	91.7	341	1	AI173687	vz77h03.r
c 874	22	91.7	331	7	CN477066	SPVDgh02	c 947	22	91.7	341	1	AV418238	AV418238
c 875	22	91.7	332	7	BB058979	BB058979	c 948	22	91.7	341	1	AA583224	nn33f08.s
c 876	22	91.7	332	5	BP103741	BP103741	c 949	22	91.7	341	2	AW484117	59293.MAR
c 877	22	91.7	332	6	CD085830	MC1-0021P	c 950	22	91.7	341	2	AW904916	RC5-NN106
c 878	22	91.7	332	6	CD085840	MC1-0021P	c 951	22	91.7	341	5	BQ975877	QHL1M14.y
c 879	22	91.7	332	6	CD085878	MC1-0021P	c 952	22	91.7	341	5	BY175860	BY175860
c 880	22	91.7	332	6	CY585028	RK029A1F0	c 953	22	91.7	341	9	CL316941	ZMMBCC050
c 881	22	91.7	332	7	CK273183	EST719261	c 954	22	91.7	342	5	BU346252	604171502
c 882	22	91.7	332	7	CN579345	MdFw2033a	c 955	22	91.7	342	5	BY176421	BY176421
c 883	22	91.7	333	2	BB100560	BB100560	c 956	22	91.7	342	7	CN281283	CN281283
c 884	22	91.7	333	2	BB116172	BB116172	c 957	22	91.7	342	8	AZ709299	RPCI-24-1
c 885	22	91.7	333	2	BF093444	CM2-TN013	c 958	22	91.7	342	8	BH337309	CH230-271

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Possible reversed clone: similarity on wrong strand Insert Length: 1988 Std Error: 0.00 Seq primer: -28mi3 rev1 ET from Amersham High quality sequence stop: 1. Location/Qualifiers

1. 58 /organism="Homo sapiens" /mol\_type="mRNA" /db\_xref="CDB:5426749" /db\_xref="taxon:9606" /clone="IMAGE:664719" /tissue\_type="neuroepithelial cells" /dev\_stage="Ntera-2 neuroepithelial cells" /lab\_host="SOLR (kanamycin resistant)" /clone\_lib="Stratagene NT2 neuronal precursor 937230" /note="Organ: brain; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Uninduced, exponentially growing neuroepithelial cells (Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGG 3' ~3' adaptor sequence: 5' CTCGAGTGTGTGTGTGTGT 3'"

FEATURES source

Alignment Scores: Pred. No.: 5.75e+03 Length: 58 Score: 22.00 Matches: 4 Percent Similarity: 80.00% Conservative: 0 Best Local Similarity: 80.00% Mismatches: 1 Query Match: 91.67% Indels: 0 Gaps: 0

US-10-030-194A-6 (1-6) x AA232607 (1-58)

Qy 1 G|VYr\*\*\*Valglu 5 ||||| ||||| ||||| |||||

Db 10 GGCTACACTGTGTGAG 24

RESULT 2

AZ389716 72 bp DNA linear GSS 02-OCT-2000

LOCUS

DEFINITION

IM0150H12R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0150H12 R, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112 USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: dunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0150 row: H column: 12

Seq primer: CACACAGGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 72. Location/Qualifiers

BH374737 AG-ND-106

AI026950 ov82909.x

AV534616 AV534616

BF610295 NXSI 057

BB153864 BB153864

BQ275246 DJ42H1.Y

BQ376620 RC4-TN018

CB100726 K156H03.Y

CN932025 000428AFB

CO329485 EK292942.

AQ822569 HS 5558.A

AL713258 DKFZP686C

AV647441 AV647441

AW064509 PICDSSH11

BE996831 NXCI 101

BY186129 BY186129

CA629896 wleln.pk0

CB706082 AMGNNUC:C

D60291 HUM097C04A

BH843105 SALK 0692

BH858609 S3\_008C.C

AW074403 xa88d08.x

BB161961 BB161961

BE702676 RC5-NN106

BF156430 f151h12.Y

BP464492 BP464492

BF752378 BF752378

F22272 HSPD06973.H

AV653900 AV653900

AV674387 AV674387

AV941939 AV941939

AW812713 RC4-ST018

BW428357 NXRV 012

BM772054 K-EST0056

BY966286 kb81a01.Y

BY790754 BY790754

CO215799 WS0042.B2

BZ126407 CH230-344

CE803128 tiger-gss-

AG211305 Oryza sat

AI468209 tg56f01.x

AA232607 58 bp mRNA linear EST 11-MAR-1998

zr28408.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone IMAGE:664719 5' similar to TR:G62337 G662337 SIMILARITY IN MIDDLE REGION OF PROTEIN TO C. ELEGANS HYPOTHETICAL PROTEIN ZK757.1.1.; mRNA sequence.

AA232607

AA232607.1 GI:1855462

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 58)

Hillier, D., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R. WashU-NCI Human EST Project

Unpublished (1997)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

RESULT 1

AA232607

LOCUS

DEFINITION

AA232607 58 bp mRNA linear EST 11-MAR-1998

zr28408.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone IMAGE:664719 5' similar to TR:G62337 G662337 SIMILARITY IN MIDDLE REGION OF PROTEIN TO C. ELEGANS HYPOTHETICAL PROTEIN ZK757.1.1.; mRNA sequence.

AA232607

AA232607.1 GI:1855462

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 58)

Hillier, D., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R. WashU-NCI Human EST Project

Unpublished (1997)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

```

source
1. .72
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCL1M0150H12"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_libs="Mouse 10kb plasmid UUGCL1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 Kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Alignment Scores:
Pred. NO.: 7.37e+03 Length: 72
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x AZ389716 (1-72)

Qy 1 GlyTyr***ValGlu 5
Db 49 GGCTACACAGTAGAA 63

RESULT 3
AL763170/c AL763170 81 bp DNA linear GSS 01-APR-2004
LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-034D07-015037,
DEFINITION genomic survey sequence.
ACCESSION AL763170
VERSION AL763170.1 GI:215111616
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoever, P. and Weissshaar, B.
TITLE GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
AUTHORS Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weissshaar, B.
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321

REFERENCE
1
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoever, P., Dekker, K.A. and Weissshaar, B.
TITLE High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
AUTHORS Li, Y., Rosso, M.G., Strizhov, N. and Weissshaar, B.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2004) Weissshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence has been recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At3g18870. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES
Location/Qualifiers
1. .81
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone_libs="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161 (GenBank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN
Alignment Scores:
Pred. NO.: 8.44e+03 Length: 81
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x AL763170 (1-81)

Qy 1 GlyTyr***ValGlu 5
Db 35 GGATACAGTGTGAA 21

RESULT 4
CG584681
LOCUS CG584681
DEFINITION OST230973 Mus musculus 129Sv/Ev Mus musculus genomic clone
ACCESSION CG584681
VERSION CG584681.1 GI:37385466
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 83)
AUTHORS Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J., Piggett, J., Beltrandeirio, H., Buxton, E.C., Edwards, J., Finch, R.A., Fiddle, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C., Key, B.W. Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D., Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z., Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A.I.
TITLE Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap

```

JOURNAL  
COMMENT  
screen to identify potential targets for therapeutic intervention  
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)  
Contact: Zambrowicz BP  
OmiBank

Lexicon Genetics Incorporated  
4000 Research Forest Drive, The Woodlands, TX 77381, USA  
Email: materials@lexgen.com  
Gene trap sequence tag generated by 3' RACE from mouse ES cells as  
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)  
Class: Gene Trap.  
Location/Qualifiers

1. .83  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129Sv/Ev"  
/db\_xref="taxon:10090"  
/clone="OST230973"  
/cell\_type="embryonic stem cell"  
/clone\_lib="Mus musculus 129Sv/Ev"

ORIGIN  
Alignment Scores:  
Pred. No.: 8.68e+03 Length: 83  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x CG584681 (1-83)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 43 GGATATGCAGTAGAA 57

RESULT 5  
AZ500158  
LOCUS  
DEFINITION  
1M0338C23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0338C23 F, genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 92)

REFERENCE  
AUTHORS  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von  
Niederhausern,A. and Wright,D.,Weiss,R.

TITLE  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

JOURNAL  
COMMENT  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0338 row: C column: 23  
Seq primer: CGTTGTAACGACGCCAGT  
Class: plasmid ends  
High quality sequence stop: 92.  
Location/Qualifiers

FEATURES  
source  
1..92  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"

/db\_xref="taxon:10090"  
/clone="UUGC1M0338C23"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

# ORIGIN

Alignment Scores:  
Pred. No.: 9.77e+03 Length: 92  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x AZ500158 (1-92)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 66 GGGTACACAGTAGAG 80

RESULT 6  
AZ773696/c

LOCUS  
DEFINITION  
2M0001N15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC2M0001N15 F, genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 92)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von  
Niederhausern,A. and Wright,D.,Weiss,R.

TITLE  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

JOURNAL  
COMMENT  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0001 row: N column: 15  
Seq primer: CGTTGTAACGACGCCAGT  
Class: plasmid ends  
High quality sequence stop: 92.

FEATURES  
source

Location/Qualifiers  
1. .92  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UNG2M0001N15"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUG1M library"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Alignment Scores:  
Pred. No.: 9.77e+03 Length: 92  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservat: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x AZ773696 (1-92)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 70 GGTATGCTGTGGAA 56

RESULT 7  
CF099395/c  
LOCUS  
DEFINITION rd72d05.y2 Meloidogyne incognita parasitic adult female SL1 TOPO v1  
Meloidogyne incognita cDNA 5', mRNA sequence.  
ACCESSION  
VERSION CF099395  
KEYWORDS  
SOURCE EST.  
ORGANISM Meloidogyne incognita (southern root-knot nematode)  
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heterodexidae; Meloidogyninae; Meloidogyne.  
1 (bases 1 to 98)  
McCarte,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,R., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarisvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
The Washington Univ. Nematode EST Project, 1999  
Unpublished (1999)  
Contact: McCarter JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by Claire Murphy and Dr. Makedonka Dautova at Washington University, St. Louis (mdautova@watson.wustl.edu). Oligo(dT)-SL1 PCR based library. Meloidogyne incognita adult cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO (Invitrogen) following the Topo TA cloning protocol. Roots infected with Meloidogyne incognita were provided by Dr. Barry Shortt at Divergence, Inc. (shortt@divergence.com) and adult females were extracted by Dr. Makedonka Dautova.

Putative full length read

The vector to vector length is 113

Seq primer: SL1 primer.

Location/Qualifiers

source

1. .98

/organism="Meloidogyne incognita"

/mol\_type="mRNA"

/db\_xref="taxon:6306"

/dev\_stage="adult females"

/lab\_host="DH108"

/clone\_lib="Meloidogyne incognita parasitic adult female SL1 TOPO v1"

/note="Vector: pCRII-TOPO (Invitrogen); Site 1: EcoRI;

Site 2: EcoRI; The library was constructed by Claire

Murphy and Dr. Makedonka Dautova at Washington University,

St. Louis. Oligo(dT)-SL1 PCR based library. Meloidogyne

incognita egg cDNA PCR products of size >400 nucleotides

containing SL1 on the 5' end and oligo(dT) on the 3' end

were non-directionally cloned into pCRII-TOPO (Invitrogen)

following the Topo TA cloning protocol. Roots infected

with Meloidogyne incognita were provided by Dr. Barry

Shortt at Divergence, Inc. and adult females were

extracted by Dr. Makedonka Dautova."

## ORIGIN

Alignment Scores:  
Pred. No.: 1.05e+04 Length: 98  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservat: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CF099395 (1-98)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 28 GGATATTCTGTGAA 14

## RESULT 8

TA367A02Q

LOCUS

DEFINITION

TA367A02Q

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

101 bp DNA linear GSS 13-DEC-2000  
T. brucei sheared genomic DNA clone 367a02, reverse sequence,  
Genomic survey sequence.

AL495214

AL495214.1 GI:11871601

GSS.

Trypanosoma brucei

Trypanosoma brucei

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

1 (bases 1 to 101)

Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,

Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,

Melville,S.E., Rajandream,M.A. and Barrell,B.G.

Direct Submission

Submitted (10-DEC-2000)

Trypanosoma brucei genome sequencing

Project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,

Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and

nhl@sanger.ac.uk

Constructed at the Institute for Genomic Research (TIGR),

Rockville, MD. Genomic DNA isolated from a cloned population of

Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).  
Email: nelsaved@tigr.org  
Details of T. brucei sequencing at the Sanger Centre are available at [http://www.sanger.ac.uk/Projects/T\\_brucei/](http://www.sanger.ac.uk/Projects/T_brucei/).

FEATURES

source  
1..101  
/organism="Trypanosoma brucei"  
/mol\_type="genomic DNA"  
/strain="TREU927"  
/db\_xref="taxon:5691"  
/clone="367a02"

ORIGIN

Alignment Scores:  
Pred. No.: 1.09e+04 Length: 101  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x TA367A02Q (1-101)

QY 1 GlyTyr\*\*\*ValGlu 5  
|||||  
DB 70 GGTACAGCGTCGAA 84

RESULT 9

AG203579/c  
LOCUS AG203579 101 bp DNA linear GSS 06-MAR-2004  
DEFINITION Pan troglodytes DNA, clone: RP43-088D13.T7, genomic survey sequence.

ACCESSION AG203579

VERSION AG203579.1 GI:45233754

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

REFERENCE

AUTHORS Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J., Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.  
BAC end sequences of Library RP-43

TITLE

JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 101)  
Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J., Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.  
Direct Submission  
TITLE Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of Bioscience and Biotechnology (KRIIB), Genome Research Center (GRC); 52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea  
JOURNAL (E-mail:redstone@mail.kribb.re.kr, URL:http://pbs.grc.kribb.re.kr/, Tel:82-42-866-7181, Fax:82-42-860-4409)  
COMMENT Clones are derived from the chimpanzee BAC library RP-43 This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: T7

LIBRARY

Vector : pBACe3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI

Location/Qualifiers

1..101

/organism="Pan troglodytes"

/mol\_type="genomic DNA"

/db\_xref="taxon:9598"

FEATURES

source

/clone="RP43-088D13.T7"  
/sex="male"  
/cell\_type="lymphocytes"  
/clone\_lib="RP-43 Chimpanzee Male BAC Library"

ORIGIN

Alignment Scores:  
Pred. No.: 1.09e+04 Length: 101  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x AG203579 (1-101)

QY 1 GlyTyr\*\*\*ValGlu 5  
|||||  
DB 74 GGCTATTCTGTAGAG 60

RESULT 10

AA249463/c  
LOCUS AA249463 105 bp mRNA linear EST 11-MAR-1997  
DEFINITION J5324.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA 5', mRNA sequence.

ACCESSION AA249463

VERSION AA249463.1 GI:1880375

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Liew,C.C.  
CDNAs from human fetal heart (1997)

TITLE

JOURNAL Unpublished (1997)

COMMENT Contact: Liew CC

Brigham and Women's Hospital

Harvard Medical School

75 Francis St. Boston, MA 02115, USA

Tel: 6177328915

Fax: 6179750995

Email: cliw@rics.bwh.harvard.edu

PCR Primers

FORWARD: 5' GCCAAGCTCGAATTAACCTCACTAAAGG 3'

BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGG 3'

Seq primer: 5' GAAATTAACCTCACTAAAGG 3'.

FEATURES

source

1..105

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/lab\_host="E. coli XL1-Blue"

/clone\_lib="Human fetal heart, Lambda ZAP Express"

/note="Vector: Lambda ZAP Express; Site 1: EcoRI; Site 2: XhoI; mRNA was purified from human fetal hearts (8-10 weeks). cDNA was synthesized using a XhoI-Oligo dt adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into predigested lambda ZAP Express."

ORIGIN

Alignment Scores:  
Pred. No.: 1.14e+04 Length: 105  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 1 Gaps: 0

US-10-030-194A-6 (1-6) x AA249463 (1-105)

QY 1 GlyTyr\*\*\*ValGlu 5



```

|||||  |||||
48 GCCTACTGTAGAA 34

RESULT 11
AA399370
LOCUS
DEFINITION
  zt57c11.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:726452
  3', mRNA sequence.
ACCESSION
  AA399370
VERSION
  AA399370.1 GI:2053115
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 106)
  Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
  Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
  Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
  White,Y., Wylie,T., Waterston,R. and Wilson,R.
  WashU-Merck EST Project 1997
  Unpublished (1997)
  Contact: Wilson RK
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available royalty-free through LNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 1135 Std Error: 0.00
  Seq primer: -41m13 fwd. ET from Amersham
  High quality sequence stop: 97.
  Location/Qualifiers
    1..106
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="GDB:5923364"
      /db_xref="taxon:9606"
      /clone="IMAGE:726452"
      /sex="male"
      /lab_host="DH10B"
      /clone_lib="Soares testis NHT"
      /note="vector: pT7T3D-Pac (Pharmacia) with a modified
  polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
  was prepared from mRNA obtained from Clontech
  Laboratories, Inc., and primed with a Not I - oligo(dT)
  primer [5'.
  TGTTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTTTTTT 3'].
  Double-stranded cDNA was ligated to Eco RI adaptors
  (Pharmacia), digested with Not I and cloned into the Not I
  and Eco RI sites of the modified pT7T3 vector. Library
  went through one round of normalization to Cot5, and was
  constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Alignment Scores:
  Pred. No.: 1..15e+04 Length: 106
  Score: 22.00 Matches: 4
  Percent Similarity: 80.00% Conservative: 0
  Best Local Similarity: 80.00% Mismatches: 1
  Query Match: 91.67% Indels: 0
  DB: 1 Gaps: 0

US-10-030-194A-6 (1-6) x AA399370 (1-106)

Qy 1 GlyTyr***ValGlu 5
|||||
Db 47 GCATATGCTGTAGAG 61

RESULT 12
BI864792
LOCUS
DEFINITION
  106 bp mRNA linear EST 11-OCT-2001
  Solanum tuberosum cDNA clone POCCR81 3' end, mRNA sequence.
ACCESSION
  CK256244
VERSION
  CK256244.1 GI:39813224
KEYWORDS
  EST.

```

```

DEFINITION
  ft96h07.y1 Zebrafish adult olfactory Danio rerio cDNA clone
  IMAGE:5283709 5', mRNA sequence.
ACCESSION
  BI864792
VERSION
  BI864792.1 GI:16057925
KEYWORDS
  EST.
SOURCE
  Danio rerio (zebrafish)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
  Cypriniformes; Cyprinidae; Danio.
REFERENCE
  1 (bases 1 to 106)
  Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
  Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
  Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
  Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
  Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
  Waterston,R. and Wilson,R.
  WashU Zebrafish EST Project 1998
  Unpublished (1998)
  Contact: Stephen L. Johnson
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: zbrafish@watson.wustl.edu
  cDNA Library Preparation: John Ngai cDNA Library Arrayed by:
  Matthew Clark. DNA Sequencing by: Washington University Genome
  Sequencing Center Clone Distribution: Genome Systems, St. Louis,
  Missouri (web address: www.genomesystems.com) (email contact:
  info@genomesystems.com) and Research Genetics, Huntsville, Alabama
  (web address: www.resgen.com) (email contact: info@resgen.com) and
  RessourcenzentrumPrimatDatenbank, Berlin, Germany (web address:
  www.rzpdp.de)
  Seq primer: T3 ET from Amersham
  High quality sequence stop: 84.
  Location/Qualifiers
    1..106
      /organism="Danio rerio"
      /mol_type="mRNA"
      /db_xref="taxon:7955"
      /clone="IMAGE:5283709"
      /sex="mixed"
      /tissue_type="Olfactory rosettes"
      /dev_stage="adult"
      /lab_host="D10Hb (Gibco BRL)"
      /clone_lib="Zebrafish adult olfactory"
      /note="vector: pSPORT1; Site 1: NotI; Site 2: SalI; This
  is a directionally cloned cDNA library from adult
  Zebrafish olfactory epithelium."

ORIGIN
Alignment Scores:
  Pred. No.: 1..15e+04 Length: 106
  Score: 22.00 Matches: 4
  Percent Similarity: 80.00% Conservative: 0
  Best Local Similarity: 80.00% Mismatches: 1
  Query Match: 91.67% Indels: 0
  DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x BI864792 (1-106)

Qy 1 GlyTyr***ValGlu 5
|||||
Db 59 GGATATGCTGTGGAA 73

RESULT 13
CK256244/c
LOCUS
DEFINITION
  106 bp mRNA linear EST 30-JUL-2004
  EST739881 potato callus cDNA library, normalized and full-length
  Solanum tuberosum cDNA clone POCCR81 3' end, mRNA sequence.
ACCESSION
  CK256244
VERSION
  CK256244.1 GI:39813224
KEYWORDS
  EST.

```

SOURCE Soluman tuberosum (potato)  
ORGANISM Soluman tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
REFERENCE Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
AUTHORS 1 (bases 1 to 106)  
TITLE Generation of ESTs from potato callus tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST739880  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr. Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: GTA ATA CGA CTC ACT ATA GGG C.  
FEATURES  
source Location/Qualifiers  
1..106  
/organism="Soluman tuberosum"  
/mol\_type="mrna"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POCCR81"  
/tissue\_type="callus"  
/lab\_host="DH10B-TonA"  
/clone\_lib="potato callus cdna library, normalized and  
full-length"  
/note="Vector: pCMVSPORT6.1; Site\_1: EcoRI; Site\_2: NotI;  
supplier: RNA was isolated from Soluman tuberosum var.  
Kennebec callus tissue grown on solid media."  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.15e+04 Length: 106  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 7 Gaps: 0  
US-10-030-194A-6 (1-6) x CK256244 (1-106)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 89 GGCTATAGTGTGAA 75  
RESULT 14  
CN027671/c linear EST 15-JUN-2004  
LOCUS UMC-p4civv1-019-a01 4-Cell-Embryo-(in vivo) p4civv Sus scrofa cdna  
DEFINITION 3', mRNA sequence.  
ACCESSION CN027671  
VERSION CN027671.1 GI:48722207  
KEYWORDS EST.  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE 1 (bases 1 to 106)  
AUTHORS Whitworth, K., Springer, G.K., Forrester, L.J., Spollen, W.G., Ries, J.,  
Lamberson, W.R., Bivens, N., Murphy, C.N., Mathialigan, N., Green, J.A.  
and Prather, R.S.  
TITLE Developmental Expression of 2,489 Gene Clusters During Pig  
Embryogenesis: An EST Project  
JOURNAL Biol. Reprod. 2 June, Epub (2004)  
COMMENT Contact: DNA Core Facility  
Animal Science - RS Prather  
University of Missouri-Columbia  
M616 Medical Sciences Bldg., Columbia, MO 65212, USA  
Tel: (573)882-0428  
Fax: (573)884-5552  
Email: bovine@rnet.missouri.edu

POLYA=Yes.

FEATURES  
source

## Location/Qualifiers

1..106  
/organism="Sus scrofa"  
/mol\_type="mrna"  
/db\_xref="taxon:9823"  
/clone\_lib="p4civv"  
/note="Funding: A grant from the Monsanto Company to the  
University of Missouri. Genetic Source:  
Slaughterhouse-derived oocytes were collected, and after  
cumulus cell removal were used for germinal vesicle stage  
oocytes, or were matured in vitro (with cumulus cells  
attached), in vitro fertilized and cultured. In vivo  
produced 4-cell and blastocyst stage embryos were  
collected on days 3 and 6, respectively. Zonae pellucidae  
were removed from the embryos prior to mRNA isolation.  
Expanded descriptions of how the tissues were collected  
can be found at the following URL:  
http://genome.rnet.missouri.edu/Swine/Methods.html.  
Library Construction (PCR Protocol): The amount of mRNA  
that was recovered from oocytes and embryos was quite  
limiting and was not sufficient for library production  
with a standard protocol. Therefore, PCR-based protocol  
was utilized for producing libraries. Poly-A RNA was  
isolated by using the MicroPoly(A) Pure kit from Ambion  
(cat. #1918). The mRNA was reverse transcribed with a  
NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide  
(Clontech) modified to contain a SalI site to generate  
full-length cDNA with a sequence complementary to the  
SMART oligonucleotide. Sequences within the SMART and dr  
oligonucleotides were used as primers to amplify the cDNAs  
by PCR with pfu turbo polymerase (Stratagene). The  
resulting PCR products were purified, digested with NotI  
and SalI and size fractionated by using a Chroma Spin-400  
followed by a Chroma Spin-1000 column (Clontech). Purified  
cDNA from each PCR reaction was ligated into the  
pCMV-SPORT6 vector. Preliminary Library Characterization:  
Randomly chosen clones from each library were analyzed by  
restriction digestion to determine average insert size (96  
clones) and by sequencing (~3 96-well plates) to confirm  
library quality [e.g. the presence of short polyA+ tails,  
genomic DNA contamination (must be <1%), ribosomal RNA  
clones (must be <1%), etc.] and to provide a sequence  
database representing the predominant clones in each  
library. The clones were sequenced at the University of  
Missouri-Columbia DNA Core Facility. Bioinformatics work  
was performed by GK Springer's bioinformatics group (WG  
Spollen, JE Ries, A Guillen) in Computer Science and  
Health Management and Informatics Departments at the  
University of Missouri-Columbia. Clone Requests: Requests  
for clones should be made to the Director of the  
University of Missouri DNA Core facility at:  
porcine@rnet.missouri.edu.  
TAG TISSUE=4-Cell-Embryo-(in vivo)  
TAG\_SEQ=GTGGC"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.15e+04 Length: 106  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 7 Gaps: 0

US-10-030-194A-6 (1-6) x CN027671 (1-106)

QY 1 GlyTyr\*\*\*ValGlu 5

Db 24 GGCTACGCCGTCGAA 10

RESULT 15

CC649660/c

LOCUS CC649660 108 bp DNA linear GSS 19-JUN-2003  
 DEFINITION OGBUD35TH\_ZM\_0.7\_1.5\_KB\_Zea\_mays\_genomic\_clone\_ZMMBMA0314E22,  
 genomic survey sequence.  
 ACCESSION CC649660  
 VERSION CC649660  
 KEYWORDS GSS, CC649660.1 GI:32052333  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 REFERENCE 1 (bases 1 to 108)  
 AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
 Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.  
 TITLE Consortium for Maize Genomics  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Cathy Whitelaw  
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.  
 FEATURES  
 source  
 1..108  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0314E22"  
 /clone\_lib="ZM 0.7 1.5 KB"  
 /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"  
 ORIGIN  
 Alignment Scores: Length: 108  
 Pred. NO.: 1.17e+04 Matches: 4  
 Score: 22.00 Conserv: 0  
 Percent Similarity: 80.00% Mismatches: 1  
 Best Local Similarity: 80.00% Indels: 0  
 Query Match: 91.67% Gaps: 0  
 DB: 9  
 US-10-030-194A-6 (1-6) x CC649660 (1-108)  
 Qy 1 GlyTyr\*\*\*ValGlu 5  
 Db 68 GGTACACAGTCGAA 54  
 RESULT 16  
 AZ309801 110 bp DNA linear GSS 29-SEP-2000  
 LOCUS 1M0017M04F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 DEFINITION clone UUGC1M0017M04 F, genomic survey sequence.  
 ACCESSION AZ309801  
 VERSION AZ309801.1 GI:10351157  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 110)  
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
 Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von  
 Niederhausern,A. and Wright,D., Weiss,R.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center

LOCUS CC649660 108 bp DNA linear GSS 19-JUN-2003  
 DEFINITION OGBUD35TH\_ZM\_0.7\_1.5\_KB\_Zea\_mays\_genomic\_clone\_ZMMBMA0314E22,  
 genomic survey sequence.  
 ACCESSION CC649660  
 VERSION CC649660  
 KEYWORDS GSS, CC649660.1 GI:32052333  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 REFERENCE 1 (bases 1 to 108)  
 AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
 Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.  
 TITLE Consortium for Maize Genomics  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Cathy Whitelaw  
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.  
 FEATURES  
 source  
 1..108  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0314E22"  
 /clone\_lib="ZM 0.7 1.5 KB"  
 /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"  
 ORIGIN  
 Alignment Scores: Length: 108  
 Pred. NO.: 1.17e+04 Matches: 4  
 Score: 22.00 Conserv: 0  
 Percent Similarity: 80.00% Mismatches: 1  
 Best Local Similarity: 80.00% Indels: 0  
 Query Match: 91.67% Gaps: 0  
 DB: 9  
 US-10-030-194A-6 (1-6) x CC649660 (1-108)  
 Qy 1 GlyTyr\*\*\*ValGlu 5  
 Db 68 GGTACACAGTCGAA 54  
 RESULT 16  
 AZ309801 110 bp DNA linear GSS 29-SEP-2000  
 LOCUS 1M0017M04F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 DEFINITION clone UUGC1M0017M04 F, genomic survey sequence.  
 ACCESSION AZ309801  
 VERSION AZ309801.1 GI:10351157  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 110)  
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
 Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von  
 Niederhausern,A. and Wright,D., Weiss,R.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center

University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0017 row: M column: 04  
 Seq primer: CGTGTAAACGACGGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 110.  
 Location/Qualifiers  
 1..110  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0017M04"  
 /sex="Male"  
 /lab\_hosts="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 [gi|4732114|gb|AF129072.1], a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

ORIGIN  
 Alignment Scores: Length: 110  
 Pred. NO.: 1.2e+04 Matches: 4  
 Score: 22.00 Conserv: 0  
 Percent Similarity: 80.00% Mismatches: 1  
 Best Local Similarity: 80.00% Indels: 0  
 Query Match: 91.67% Gaps: 0  
 DB: 8  
 US-10-030-194A-6 (1-6) x AZ309801 (1-110)  
 Qy 1 GlyTyr\*\*\*ValGlu 5  
 Db 34 GGTACACAGTCGAA 20  
 RESULT 17  
 AI664163 114 bp mRNA linear EST 10-MAY-1999  
 LOCUS ue79g10.r1 Soares\_NMPu Mus musculus CDNA clone IMAGE:1497378.5  
 DEFINITION similar to gb:L25080 TRANSFORMING PROTEIN RHOA (HUMAN); mRNA  
 sequence.  
 ACCESSION AI664163  
 VERSION AI664163.1 GI:4767746  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 114)  
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL COMMENT  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:934982  
Trace considered overall poor quality  
Putative full length read  
vector to vector length is 215  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 1.  
FEATURES  
source  
1. .114  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1497378"  
/sex="female"  
/dev\_stages="adult"  
/lab\_host="DH10B"  
/clone\_lib="Soares\_NMPu"  
/note="Organ: uterus; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pregnant mouse uterus, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.25e+04 Length: 114  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 1 Gaps: 0  
US-10-030-194A-6 (1-6) x AI664163 (1-114)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 73 GGATACAGTGTGGAG 59  
RESULT 18  
CA353391 116 bp mRNA linear EST 05-NOV-2002  
LOCUS 624896 CCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT69P17\_B\_H09 5', mRNA sequence.  
DEFINITION CA353391  
VERSION CA353391.1 GI:24598562  
KEYWORDS EST.  
SOURCE Oncorhynchus mykiss (rainbow trout)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
REFERENCE 1 (bases 1 to 116)  
AUTHORS Rextroad,C.E. 3rd, Lee,Y., Keele,J.W., Karamycheva,S., Brown,G., Koop,B., Gahr,S.A., Palti,Y. and Quackenbush,J.  
TITLE Sequence analysis of a rainbow trout cDNA library and creation of a gene index  
JOURNAL CytoGenet. Genome Res. 102 (1-4), 347-354 (2003)  
COMMENT Contact: Rextroad CE  
USDA, ARS, National Center for Cool and Cold Water Aquaculture  
11876 Leetown Road, Kearneysville, WV 25430, USA  
Tel: 304 724 8340 x2129  
Fax: 304 725 0351  
Email: crexroad@cccwa.ars.usda.gov  
Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim\_alt option. Vector identified by cross\_match v0.990329.

Seq primer: AGCGGATAACAAATTTCACACAGGA.  
FEATURES  
source  
1. .116  
Location/Qualifiers  
/organism="Oncorhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/clone="1RT69P17\_B\_H09"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="NCCCWA 1RT"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; library made from pooled tissue from brain, gill, liver, spleen, muscle, and kidney."  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.27e+04 Length: 116  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 6 Gaps: 0  
US-10-030-194A-6 (1-6) x CA353391 (1-116)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 11 GGATACAGTGTGGAA 25  
RESULT 19  
AI906955/c 118 bp mRNA linear EST 30-MAR-2000  
LOCUS OV-BT130-170399-007 BT130 Homo sapiens cDNA, mRNA sequence.  
DEFINITION AI906955  
ACCESSION AI906955  
VERSION AI906955.1 GI:6497369  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 118)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/seq/gethtml.pl?t1=QV&t2=QV-BT130-007.html&t3=170399&t4=1)  
Seq primer: puc 18 forward.  
FEATURES  
source  
1. .118  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="female"  
/dev\_stages="Adult"  
/clone\_lib="BT130"

/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

## Alignment Scores:

Pred. No.:	1.3e+04	Length:	118
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	1	Gaps:	0

US-10-030-194A-6 (1-6) x AI906955 (1-118)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 107 GCGTATCCGTGGAG 93

## RESULT 20

BUB16086/c

## LOCUS

DEFINITION BUB16086 118 bp mRNA linear EST 15-OCT-2002  
 N060C10 Populus bark cDNA library Populus tremula x Populus tremuloides cDNA 5 prime, mRNA sequence.

## ACCESSION

VERSION BUB16086

## KEYWORDS

EST.

## ORGANISM

SOURCE Populus tremula x Populus tremuloides  
 ORGANISM Populus tremula x Populus tremuloides  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosids; eurosids I; Malpighiales; Salicaceae; Populus.

## REFERENCE

AUTHORS 1 (bases 1 to 118)

TITLE Unneberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.

COMMENT The poplar tree transcriptome: Analysis of expressed sequence tags from multiple libraries

JOURNAL Unpublished (2002)

CONTACT: BHALERAO RUPALI R.

Umea Plant Science Center

Department of Plant Physiology

University of Umea, 901 87 Umea, Sweden

Tel: +46 90 786 5279

Fax: +46 90 786 6676

Email: rupali.bhalerao@plantphys.umu.se.

## FEATURES

source

1..118

/organism="Populus tremula x Populus tremuloides"

/mol\_type="mRNA"

/db\_xref="taxon:47664"

/tissue\_type="bark"

/clone\_lib="Populus bark cDNA library"

## ORIGIN

## Alignment Scores:

Pred. No.:	1.3e+04	Length:	118
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	5	Gaps:	0

US-10-030-194A-6 (1-6) x BUB16086 (1-118)

## Qy

1 GlyTyr\*\*\*ValGlu 5

|||||

Db 58 GGGTATACCGTGGAG 44

## RESULT 21

AW935423

## LOCUS

DEFINITION CM3-DT0004-110500-179-a04 DT0004 Homo sapiens cDNA, mRNA sequence.

## ACCESSION

AW935423

## VERSION

AW935423.1

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 126)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Masukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=st2=CM3-DT0004-110

500-179-a04&t3=2000-05-11&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 126.

## FEATURES

source

1..126

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev stage="Adult"

/clone\_lib="DT0004"

/note="Organ: denis\_drash; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

## Alignment Scores:

Pred. No.:	1.4e+04	Length:	126
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	2	Gaps:	0

## ORIGIN

US-10-030-194A-6 (1-6) x AW935423 (1-126)

## Qy

1 GlyTyr\*\*\*ValGlu 5

|||||

Db 7 GGGTATGCTGTCGAA 21

## RESULT 22

AZ083244

## LOCUS

DEFINITION RPCI-23-467A14.TJB RPCI-23 Mus musculus genomic clone

127 bp DNA linear GSS 08-MAY-2000

ACCESSION RPCI-23-467A14, genomic survey sequence.

AW935423

## KEYWORDS

GSS.

```

SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 127)
Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Kroi, M., de
Jong, P. and Frazer, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Contact: Shaving Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 467 row: A column: 14
Seq primer: SP6
Classes: BAC ends.
FEATURES             Location/Qualifiers
     source           1..127
     /organism="Mus musculus"
     /mol_type="genomic DNA"
     /strain="C57BL/6J"
     /db_xref="taxon:10090"
     /clone="RPCI-23-467A14"
     /sex="Female"
     /lab_host="DH10B"
     /clone_lib="RPCI-23"
     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies). "
```

```

ORIGIN
Alignment Scores:
Pred. No.: 1.43e+04 Length: 128
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: Gaps: 0

US-10-030-194A-6 (1-6) x AJ770527 (1-128)

Qy 1 GYTYR***ValGlu 5
|||
Db 37 GGATATGCTGTGGAA 51

RESULT 24
BG994255/c 129 bp mRNA linear EST 13-JUN-2001
LOCUS CM3-HT1150-130201-741-e03 HT1150 Homo sapiens cDNA, mRNA sequence.
DEFINITION BG994255
ACCESSION BG994255.1 GI:14398325
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 129)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
PUBMED
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-HT1150-
130201-741-e03&t3=2001-02-13&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 46
High quality sequence stop: 129.

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
```

```

SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 127)
Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Kroi, M., de
Jong, P. and Frazer, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Contact: Shaving Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 467 row: A column: 14
Seq primer: SP6
Classes: BAC ends.
FEATURES             Location/Qualifiers
     source           1..127
     /organism="Mus musculus"
     /mol_type="genomic DNA"
     /strain="C57BL/6J"
     /db_xref="taxon:10090"
     /clone="RPCI-23-467A14"
     /sex="Female"
     /lab_host="DH10B"
     /clone_lib="RPCI-23"
     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies). "
```

```

FEATURES
  source
    Location/Qualifiers
      1..129
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /dev_stage="Adult"
        /clone_lib="HTL150"
        /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
        Site 2: SmaI; A mini-library was made by cloning products
        derived from ORSTES PCR (U.S. Letters Patent application
        No. 196,716 - Ludwig Institute for Cancer Research)
        profiles into the pUC 18 vector. Reverse transcription of
        tissue mRNA and cDNA amplification were performed under
        low stringency conditions."
ORIGIN
Alignment Scores:
Pred. No.: 1.44e+04 Length: 129
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0
US-10-030-194A-6 (1-6) x BG994255 (1-129)

Qy 1 GlyTyr***ValGlu 5
Db 92 GGCTACAGTGTGGAG 78

RESULT 25
CC356105/c
LOCUS 9712 Medical Center Drive, Rockville, MD 20850, USA
DEFINITION Zea mays genomic clone ZMMBT506N18,
genomic survey sequence.
ACCESSION CC356105
VERSION CC356105.1 GI:30825505
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 129)
AUTHORS Whetlaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Reinick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: FUIPV817B
Contact: Cathy Whitelaw
TIGR
7912 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
  source
    Location/Qualifiers
      1..129
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMMBT506N18"
        /clone_lib="ZM_0.6-1.0 KB"
        /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
        Cot selected genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.: 1.44e+04 Length: 129
Score: 22.00 Matches: 4

```

```

Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 8 Gaps: 0
US-10-030-194A-6 (1-6) x CC356105 (1-129)

Qy 1 GlyTyr***ValGlu 5
Db 58 GGGTATTTCAGTTGAA 44

RESULT 26
AL766514/c
LOCUS AL766514 129 bp DNA linear GSS 01-APR-2004
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-211D06-014132,
genomic survey sequence.
ACCESSION AL766514
VERSION AL766514.1 GI:21519647
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weishaar,B.
GABI-Kat Simplesearch: a flanking sequence tag (FST) database for
the identification of T-DNA insertion mutants in Arabidopsis
thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
REFERENCE 2
AUTHORS Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
Weishaar,B.
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321
REFERENCE 3
AUTHORS Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
Weishaar,B.
High-throughput generation of sequence indexes from T-DNA
mutagenized Arabidopsis thaliana lines
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE 4 (bases 1 to 129)
AUTHORS Li,Y., Strizhov,N., Rosso,M.G. and Weishaar,B.
Direct Submission
Submitted (31-MAR-2004) Weishaar B., Max-Planck-Institut fuer
Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence has been recovered from the left border of the T-DNA.
It indicates an insertion close to or within gene At3g18870.
Details on the protocols used for generation of the sequence are
described in References 1-3. The sequences are generated at the MPI
for Plant Breeding Research in the context of the GABI-Kat project.
GABI-Kat is part of the German Plant Genomics program designated
'GABI'. Information on line availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.
Location/Qualifiers
  1..129
    /organism="Arabidopsis thaliana"
    /mol_type="genomic DNA"
    /strain="Columbia 0"
    /db_xref="taxon:3702"
    /clone="GK-211D06-014132"
    /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
    /ecotype="Col-0"
    /note="PCR was performed on DNA from Arabidopsis thaliana
    plants (T1) which were transformed with the T-DNA from
    vector pAC161 (GenBank accession number: AJ537314). The
    lines contain one or more T-DNA insertions. The DNA

```

fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

```

ORIGIN
Alignment Scores:
Pred. No.: 1.44e+04 Length: 129
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x AL766514 (1-129)

QY 1 GlyTyr***ValGlu 5
DB 98 GGATACAGTGTGAA 84

RESULT 27
CG644531 129 bp DNA linear GSS 02-OCT-2003
LOCUS OST386390 Mus musculus 129Sv/Ev Mus musculus genomic clone
DEFINITION OST386390, genomic survey sequence.
ACCESSION CG644531
VERSION CG644531.1 GI:37468380
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 129)
AUTHORS Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
Key,B.W. Jr., Kipp,P., Kohlhaufl,B., Ma,Z.-Q., Markesich,D.,
Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
Zhu,Q., Person,C. and Sands,A.T.
TITLE Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT Contact: Zambrowicz BP
Omnibank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.
FEATURES
Location/Qualifiers
source
1..129
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
/clone="OST386390"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

ORIGIN
Alignment Scores:
Pred. No.: 1.44e+04 Length: 129
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x CG644531 (1-129)

QY 1 GlyTyr***ValGlu 5
DB 98 GGATACAGTGTGAA 84

```

```

Db 71 GGTACAGTGTGAG 85

RESULT 28
BG140779 130 bp mRNA linear EST 31-JAN-2001
LOCUS EST481221 wild tomato pollen Lycopersicon pennellii cDNA clone
DEFINITION GUPPI8K10 5', sequence, mRNA sequence.
ACCESSION BG140779
VERSION BG140779.1 GI:12640968
KEYWORDS EST.
SOURCE Lycopersicon pennellii
ORGANISM Lycopersicon pennellii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 130)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from wild tomato (L. pennellii) pollen
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
Location/Qualifiers
source
1..130
/organism="Lycopersicon pennellii"
/mol_type="mRNA"
/cultivar="TA56"
/db_xref="taxon:28526"
/clone="GUPPI8K10"
/tissue_type="pollen"
/dev_stage="pollen collected from open flowers"
/lab_host="SOLR"
/clone_lib="wild tomato pollen"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Pollen was collected from open flowers from
L.pennellii TA56, and stored at -80 C until library
construction."
ORIGIN
Alignment Scores:
Pred. No.: 1.45e+04 Length: 130
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x BG140779 (1-130)

QY 1 GlyTyr***ValGlu 5
DB 116 GGCTACTACTGTGGAG 130

RESULT 29
BG385119 130 bp mRNA linear EST 12-MAR-2001
LOCUS BG385119 306732 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BG385119
ACCESSION BG385119.1 GI:13309591
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 130)
AUTHORS Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,
Quackenbush,J. and Keele,J.W.
TITLE Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly

```



JOURNAL MEDLINE PUBMED COMMENT

Mamm. Genome 13 (8), 475-478 (2002)  
22213789  
12226715

Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390

Email: smith@mail.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.

PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCACGTCACGACG  
Plate: 94 row: G column: 5  
Seq primer: ATTGAGTGACACTATAG.

FEATURES source  
1..130  
Location/Qualifiers  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="MARC 1P1G"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from day 11, 13, 15, 20,  
and 30 embryos."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.45e+04 Length: 130  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservatived: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x BG385119 (1-130)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 4 GTTACACCGTGGAA 18

RESULT 30  
AW750597

LOCUS AW750597 131 bp mRNA linear EST 28-APR-2000

DEFINITION PM3-CN0029-190100-001-e11 CN0029 Homo sapiens cDNA, mRNA sequence.

ACCESSION AW750597.1 GI:7665529

VERSION AW750597.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

CONTACT: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922

Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM3&t2=PM3-CN0029-190100-001-e11&t3=2000-01-19&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 25  
High quality sequence stop: 42.  
Location/Qualifiers  
1..131  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="CN0029"  
/note="Organ: colon normal; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.47e+04 Length: 131  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservatived: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x AW750597 (1-131)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 30 GGCTACACCGTAGAG 44

RESULT 31  
BQ976013

LOCUS BQ976013 131 bp mRNA linear EST 21-AUG-2002

DEFINITION QH120D04.yg.ab1 QH\_ABCDI sunflower RH801 Helianthus annuus cDNA clone QH120D04, mRNA sequence.

ACCESSION BQ976013

VERSION BQ976013.1 GI:22393536

KEYWORDS EST.

SOURCE Helianthus annuus (common sunflower)

ORGANISM Helianthus annuus

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.

AUTHORS 1 (bases 1 to 131)  
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.

TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project

JOURNAL http://compnogens.ucdavis.edu/unpublished (2002)

COMMENT Contact: Alexander Kozik [R.W.Michelmore] Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Amundson Hall, UCD, Davis, CA 95616, USA Tel: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu] belongs to contig QH\_CA Contig2725, see http://cgdb.ucdavis.edu/ for details.  
Plate: QH120 row: D column: 04.  
Location/Qualifiers  
1..131

```

/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QH120D04"
/lab_host="E.coli"
/clone_lib="QH_ABCDI sunflower RHA801"
/note="Vector: pBRCNASFIAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgdb.ucdavis.edu/
TAG_TISSUE=flowers environmental stress
TAG_LIB=QH ABCDI sunflower RHA801
TAG_SEQ=CGATGCGGG"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 1.47e+04 Length: 131
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 5 Gaps: 0

```

US-10-030-194A-6 (1-6) x BQ976013 (1-131)

```

QY 1 GYTYR***ValGlu 5
|||||
Db 1 GGGTACAGTGTGGAA 15

```

## RESULT 32

```

BQ978589
LOCUS BQ978589 131 bp mRNA linear EST 21-AUG-2002
DEFINITION clone QH15H23, mRNA sequence.

```

```

ACCESSION BQ978589
VERSION BQ978589.1 GI:22396112
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)

```

## ORGANISM

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.

```

```

REFERENCE 1 (bases 1 to 131)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,B. and Bradford,K.
TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compenomics.ucdavis.edu/
JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmil.ucdavis.edu]
belongs to contig QH_CA_Contig2725, see http://cgdb.ucdavis.edu/
for details.
Plate: QH15 row: H column: 23.

```

## JOURNAL

## COMMENT

```

The library was characterised by oligonucleotide fingerprinting
(ONF) to reduce sequencing redundancy. According to the ONF
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONF cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well as the coordinates of
the other clones assigned to the same ONF cluster as the clone from
which the above EST is generated is available at the sea urchin
project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA
clones and filters are distributed via the Resource Center/Primary
Database of the German Human Genome Project (http://www.rzpd.de)
PCR Primers
FORWARD: 5' CCCACGGCTTTACACTTTATGCTTCGCGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTACGCAGCTGCGGAAGGGGATGTC 3' (M13FSP) 3'-seq

```

## FEATURES

## source

```

1. .131
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"

```

```

/clone="QH15H23"
/lab_host="E.coli"
/clone_lib="QH ABCDI sunflower RHA801"
/note="Vector: pBRCNASFIAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgdb.ucdavis.edu/
TAG_TISSUE=flowers environmental stress
TAG_LIB=QH ABCDI sunflower RHA801
TAG_SEQ=CGATGCGGG"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 1.47e+04 Length: 131
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 5 Gaps: 0

```

US-10-030-194A-6 (1-6) x BQ978589 (1-131)

```

QY 1 GYTYR***ValGlu 5
|||||
Db 1 GGGTACAGTGTGGAA 15

```

## RESULT 33

```

CD297327
LOCUS CD297327 131 bp mRNA linear EST 16-SEP-2003
DEFINITION StrPu691.008722 Sea urchin larva cDNA library MPWGp691
Strongylocentrotus purpuratus cDNA clone
MPWGp691G2415;MPI_SURUDI_15G24 5', mRNA sequence.

```

```

ACCESSION CD297327
VERSION CD297327.1 GI:34748404
KEYWORDS EST.
SOURCE Strongylocentrotus purpuratus

```

## ORGANISM

```

Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.

```

```

1 (bases 1 to 131)
Poustka,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,
Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.

```

```

Generation, annotation, evolutionary analysis, and database
integration of 20,000 unique sea urchin EST clusters

```

```

Genome Res. 13 (12), 2736-2746 (2003)

```

```

Contact: Poustka AJ,
Laboraty 145, dept.Lehrach
Max-Planck-Institut fuer Molekulare Genetik

```

```

Innestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: poustka@molgen.mpg.de

```

```

The library was characterised by oligonucleotide fingerprinting
(ONF) to reduce sequencing redundancy. According to the ONF
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONF cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well as the coordinates of
the other clones assigned to the same ONF cluster as the clone from
which the above EST is generated is available at the sea urchin
project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA
clones and filters are distributed via the Resource Center/Primary
Database of the German Human Genome Project (http://www.rzpd.de)
PCR Primers
FORWARD: 5' CCCACGGCTTTACACTTTATGCTTCGCGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTACGCAGCTGCGGAAGGGGATGTC 3' (M13FSP) 3'-seq

```

Seq primer: 5'-CCGTCGCCGAATCCGGGT-3' pSport3/86

High quality sequence stop: 131.

## FEATURES

source  
1. .131  
Location/Qualifiers  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="MPMGp691G2415;MPI SURUDI\_15G24"  
/tissue\_type="whole larva"  
/dev\_stage="larva 2-3 weeks"  
/lab\_host="E.coli, XL1 blue"  
/clone\_lib="Sea urchin larva cDNA library MPMGp691"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI; Random  
primed and directionally cloned in pSport1 vector using a  
NotI (5'-pGACTGATCTAGATCGGCGCGGCC (T)15-3' and a  
SalI 5'-TCGACCCGCGCTCCG-3' adapters (Gibco BRL)"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.47e+04 Length: 131  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservatative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CD297327 (1-131)

QY 1 GlyTyr\*\*\*ValGlu 5

Db 53 GGATATCTGTAGAG 67

## RESULT 34

CF088605

LOCUS CF088605 131 bp mRNA linear EST 22-JUL-2003  
DEFINITION QHM1F08.yg.abl QH M sunflower H. argophyllus Helianthus argophyllus  
cDNA clone QHM1F08, mRNA sequence.

ACCESSION CF088605

VERSION CF088605.1 GI:33127672

KEYWORDS EST.

SOURCE

ORGANISM

Helianthus argophyllus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;  
asterids; campanulids; Asterales; Asteraceae; Asteroideae;  
Heliantheae; Helianthus.

1 (bases 1 to 131)

REFERENCE  
AUTHORS  
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,  
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,  
Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,  
Lai, Z., Church, S., Jackson, L. and Bradford, K.

Lettuce and Sunflower ESTs from the Compositae Genome Project

http://compgenomics.ucdavis.edu/

Unpublished (2002)

## JOURNAL

COMMENT

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Asmundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]

belongs to contig QH\_CA\_Contig4345, see http://cgdb.ucdavis.edu/

for details.

Plate: QHM1 row: F column: 08.

## FEATURES

source

1. .131  
Location/Qualifiers

/organism="Helianthus argophyllus"

/mol\_type="mRNA"

/db\_xref="taxon:73275"

/clone="QHMF08"

/lab\_host="E.coli"

/clone\_lib="QH M sunflower H. argophyllus"

/note="Vector: pBRcDNAsfiAB; The library was constructed

## ORIGIN

Alignment Scores:

Pred. No.: 1.47e+04 Length: 131  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservatative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CF088605 (1-131)

QY 1 GlyTyr\*\*\*ValGlu 5

Db 3 GGCTATCTGTGGAG 17

## RESULT 35

AZ752194/c

LOCUS AZ752194 131 bp DNA linear GSS 25-JAN-2001  
DEFINITION RPCI-24-147L16.TJ RPCI-24 Mus musculus genomic clone  
RPCI-24-147L16, genomic survey sequence.

ACCESSION AZ752194

VERSION AZ752194.1 GI:12537452

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 131)

REFERENCE

AUTHORS

Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinet, B., Levins, M.,  
Teegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorgis, E.,  
Russell, D., de Jong, P. and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-24  
Unpublished (1999)

TITLE

JOURNAL

COMMENT

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC

library availability, please contact Pieter de Jong

(pdejong@mail.cho.org). Clones may be purchased from BACPAC

Resources (<http://www.choi.org/bacpac/orderingframe.html>). BAC end

page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)

Plate: 147 row: L column: 16

Seq primer: SP6

Class: BAC ends.

## FEATURES

source

1. .131  
Location/Qualifiers

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="RPCI-24-147L16"

/sex="Male"

/cell\_type="Spleen/Brain"

/clone\_lib="RPCI-24"

/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI;

RPCI-24 Mouse BAC Library produced by Pieter de Jong. The

library was cloned in the pTARBAC1 cloning vector at the

BamHI sites using MboI partially digested male C57BL/6J

DNA."

## ORIGIN

Alignment Scores:

from three different sources (seedling, root and leaf) of  
RNA from a single genotype. cDNAs were pooled and  
directionally cloned into a custom medium-copy vector.  
Details of library construction can be obtained at  
<http://cgdb.ucdavis.edu/>

Pred. No.: 1.47e+04 Length: 131  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x AZ752194 (1-131)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 ||||| |||||  
 Db 112 GGCTACTCTGTGGAG 98

RESULT 36

LOCUS BF557111/c 132 bp mRNA linear EST 12-DEC-2000  
 DEFINITION UI-R-CO-gr-d-01-0-UI.r1 UI-R-CO Rattus norvegicus cDNA clone  
 UI-R-CO-gr-d-01-0-UI 5', mRNA sequence.

ACCESSION BF557111.1 GI:11666841

VERSION EST.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus

1 (bases 1 to 132)

REFERENCE Bonaldo,M.F., Lennon,G. and Soares,M.B.

AUTHORS Normalization and subtraction: two approaches to facilitate gene  
 TITLE discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704477

PUBMED 889548

COMMENT

Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
 clones will be available through Research Genetics (www.resgen.com)

This clone is also available through the I.M.A.G.E. Consortium at  
 LLNL (info@image.llnl.gov). IMAGE ID= 1772818

Seq primer: M13 Forward

FEATURES

source

1..132  
 Location/Qualifiers  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-CO-gr-d-01-0-UI"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="UI-R-CO"

/note="Vector: pTT3D-Pac (Pharmacia) with a modified  
 polylinker; Site1: Not I; Site 2: Eco RI; The UI-R-CO  
 library is a subtracted library derived from the UI-R-Al  
 and UI-R-E1 libraries. The UI-R-Al library consisted of a  
 mixture of individually tagged normalized libraries  
 constructed from rat placenta, adult lung, brain, liver,  
 kidney, heart, spleen, ovary, and muscle. The UI-R-E1  
 library consisted of a mixture of individually tagged  
 normalized libraries constructed from 8, 12 and 18-day  
 embryo. The tag is a string of 3-5 nucleotides present  
 between the Not I site and the oligo-dT track which  
 allows identification of the library of origin of a clone  
 within the mixture. The subtracted library (UI-R-CO) was  
 constructed as follows: PCR amplified cDNA inserts from a  
 pool of UI-R-Al and UI-R-E1 clones from which 3' ESTs had  
 been derived was used as a driver in a hybridization with  
 the pooled UI-R-Al and UI-R-E1 library in the form of  
 single-stranded circles. The remaining single-stranded

circles (subtracted library) was purified by  
 hydroxyapatite column chromatography, converted to  
 double-stranded circles and electroporated into DH10B  
 bacteria (Life Technologies) to generate the UI-R-CO  
 library. This procedure has been previously described  
 (Bonaldo, Lennon and Soares, Genome Research 6: 791-806,  
 1996)"

ORIGIN

Alignment Scores: 1.48e+04 Length: 132  
 Pred. No.: 22.00 Matches: 4  
 Score: 80.00% Conservative: 0  
 Percent Similarity: 80.00% Mismatches: 1  
 Best Local Similarity: 91.67% Indels: 0  
 Query Match: 91.67% Gaps: 0  
 DB: 2

US-10-030-194A-6 (1-6) x BF557111 (1-132)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 ||||| |||||  
 Db 22 GGCTATCGGTGCGAG 8

RESULT 37

LOCUS BI843879

DEFINITION BI843879 132 bp mRNA linear EST 04-OCT-2001  
 ft91e03.y1 Zebrafish adult olfactory Danio rerio cDNA clone  
 IMAGE:5283172 5', mRNA sequence.

ACCESSION BI843879.1 GI:15956402

VERSION EST.

KEYWORDS Danio rerio (zebrafish)

SOURCE

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 132)

REFERENCE Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,  
 Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,

Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,  
 Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,

Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,  
 Waterston,R. and Wilson,R.

WashU Zebrafish EST Project 1998

Unpublished (1998)

Other ESTs: ft91e03.x1

Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu

cDNA Library Preparation: John Ngai cDNA Library Arrayed by:  
 Mathew Clark. DNA Sequencing by: Washington University Genome

Sequencing Center Clone distribution: Genome Systems, St. Louis,  
 Missouri (web address: www.genomesystems.com) (email contact:

info@genomesystems.com) and Research Genetics, Huntsville, Alabama  
 (web address: www.resgen.com) (email contact: info@resgen.com) and

RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:

www.rzpd.de)

Seq primer: T3 ET from Amersham

High quality sequence stop: 131.

Location/Qualifiers

source

1..132  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7955"  
 /clone="IMAGE:5283172"  
 /sex="mixed"  
 /tissue\_type="Olfactory rosettes"  
 /dev\_stage="adult"  
 /lab\_host="D10HB (Gibco BRL)"  
 /clone\_lib="Zebrafish adult olfactory"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This is a directionally cloned cDNA library from adult zebrafish olfactory epithelium."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.48e+04 Length: 132  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservativeness: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x BI843879 (1-132)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 28 GGATATGCTGTGGAA 42

## RESULT 38

BU020428

LOCUS

DEFINITION BU020428 132 bp mRNA linear EST 23-AUG-2002  
 clone QHE27G13.YG.ab1 OH\_EFGHJ sunflower RHA280 Helianthus annuus cDNA

ACCESSION BU020428.1 GI:22455948

VERSION EST.

KEYWORDS Helianthus annuus (common sunflower)

SOURCE

ORGANISM

REFERENCE  
 AUTHORS  
 Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damm, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.

## TITLE

Lettuce and Sunflower ESTs from the Compositae Genome Project

http://compgenomics.ucdavis.edu/

## JOURNAL

COMMENT

Contact: Alexander Kozik [R.W.Michelmore]  
 Department of Vegetable Crops, R.W.Michelmore Lab  
 University of California at Davis (UCD)  
 Asmundo Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]  
 Belongs to contig QH\_CA\_Contig2725, see http://cgdb.ucdavis.edu/  
 for details.  
 Plate: QHE27 row: G column: 13.

## FEATURES

source

1..132 Location/Qualifiers

/organism="Helianthus annuus"

/mol\_type="mRNA"

/cultivar="RHA280"

/db\_xref="taxon:4232"

/clone="QHE27G13"

/lab\_host="E. coli"

/clone\_lib="QH\_EFGHJ sunflower RHA280"

/note="Vector: pBRCNASTAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/

TAG TISSUE=hulls

TAG\_LIB=QH\_EFGHJ sunflower RHA280

TAG\_SEQ=CCGCTAGTCGGG"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.48e+04 Length: 132  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservativeness: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 5 Gaps: 0

US-10-030-194A-6 (1-6) x BU020428 (1-132)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 1 GGGTACAGTGTGGAA 15

## RESULT 39

CG653419

LOCUS

DEFINITION CG653419 132 bp DNA linear GSS 02-OCT-2003  
 OST418752 Mus musculus 129Sv/Ev Mus musculus genomic clone

ACCESSION CG653419

VERSION GSS.

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM

REFERENCE  
 AUTHORS  
 Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J., Piggott, J., BeltrandelRio, H., Buxton, E.C., Edwards, J., Finch, R.A., Friddle, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C., Key, B.W., Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D., Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z., Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A.T.

Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention

Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

Contact: Zambrowicz BP

Omnibank

Lexicon Genetics Incorporated

4000 Research Forest Drive, The Woodlands, TX 77381, USA

Email: materials@lexgen.com

Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)

Class: Gene Trap.

## FEATURES

source

1..132 Location/Qualifiers

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="129Sv/Ev"

/db\_xref="taxon:10090"

/clone="OST418752"

/cell\_type="embryonic stem cell"

/clone\_lib="Mus musculus 129Sv/Ev"

## ORIGIN

Alignment Scores:

Pred. No.: 1.48e+04 Length: 132  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservativeness: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x CG653419 (1-132)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 14 GGCTACTCAGTAGAA 28

## RESULT 40

AU009303

LOCUS AU009303 134 bp mRNA linear EST 31-JUL-1998  
 DEFINITION AU009303 Schizosaccharomyces pombe late log phase cDNA  
 Schizosaccharomyces pombe cDNA clone spc04757, mRNA sequence.  
 ACCESSION AU009303  
 VERSION AU009303.1 GI:3345983  
 KEYWORDS EST.  
 SOURCE Schizosaccharomyces pombe (fission yeast)  
 ORGANISM Schizosaccharomyces pombe  
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 Schizosaccharomycetales; Schizosaccharomycetaceae;  
 Schizosaccharomycetes.  
 REFERENCE 1 (bases 1 to 134).  
 AUTHORS Morimyo,M. and Mita,K.  
 TITLE Identification of expressed sequence tags of Schizosaccharomyces pombe  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Mitsuoki Morimyo  
 Genome Research Group  
 National Institute of Radiological Sciences  
 9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan  
 Email: morimyo@nirs.go.jp.  
 Location/Qualifiers  
 1. .134  
 /organism="Schizosaccharomyces pombe"  
 /mol\_type="mRNA"  
 /strain="972"  
 /db\_xref="taxon:4896"  
 /clone="spc04757"  
 /sex="h minus"  
 /note="Vector: M13mp19; The cDNA library of Schizosaccharomyces pombe was prepared by cloning cDNA into the SmaI site of M13mp19 DNA and the direction of DNA sequences was not always from 5' to 3'. The cDNA data of Schizosaccharomyces pombe are available for searching on the World Wide Web. (URL, http://www.nirs.go.jp)"  
 US-10-030-194A-6 (1-6) x AU009303 (1-134)  
 QY 1 GlyTyr\*\*\*ValGlu 5  
 ||||| |||||  
 Db 16 GGATATTTCAGTAGAA 30  
 RESULT 41  
 EQ293027  
 LOCUS AU009303 134 bp mRNA linear EST 31-JUL-1998  
 DEFINITION AU009304 Schizosaccharomyces pombe late log phase cDNA  
 Schizosaccharomyces pombe cDNA clone spc04760, mRNA sequence.  
 ACCESSION AU009304  
 VERSION AU009304.1 GI:3345984  
 KEYWORDS EST.  
 SOURCE Schizosaccharomyces pombe (fission yeast)  
 ORGANISM Schizosaccharomyces pombe  
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 Schizosaccharomycetales; Schizosaccharomycetaceae;  
 Schizosaccharomycetes.  
 REFERENCE 1 (bases 1 to 134).  
 AUTHORS Morimyo,M. and Mita,K.  
 TITLE Identification of expressed sequence tags of Schizosaccharomyces pombe  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Mitsuoki Morimyo  
 Genome Research Group  
 National Institute of Radiological Sciences  
 9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan  
 Email: morimyo@nirs.go.jp.  
 Location/Qualifiers  
 1. .134  
 /organism="Schizosaccharomyces pombe"  
 /mol\_type="mRNA"  
 /strain="972"  
 /db\_xref="taxon:4896"  
 /clone="spc04757"  
 /sex="h minus"  
 /note="Vector: M13mp19; The cDNA library of Schizosaccharomyces pombe was prepared by cloning cDNA into the SmaI site of M13mp19 DNA and the direction of DNA sequences was not always from 5' to 3'. The cDNA data of Schizosaccharomyces pombe are available for searching on the World Wide Web. (URL, http://www.nirs.go.jp)"  
 US-10-030-194A-6 (1-6) x AU009303 (1-134)  
 QY 1 GlyTyr\*\*\*ValGlu 5  
 ||||| |||||  
 Db 16 GGATATTTCAGTAGAA 30  
 RESULT 41  
 EQ293027  
 LOCUS AU009304 134 bp mRNA linear EST 31-JUL-1998  
 DEFINITION AU009304 Schizosaccharomyces pombe late log phase cDNA  
 Schizosaccharomyces pombe cDNA clone spc04760, mRNA sequence.  
 ACCESSION AU009304  
 VERSION AU009304.1 GI:3345984  
 KEYWORDS EST.  
 SOURCE Schizosaccharomyces pombe (fission yeast)  
 ORGANISM Schizosaccharomyces pombe  
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 Schizosaccharomycetales; Schizosaccharomycetaceae;  
 Schizosaccharomycetes.  
 REFERENCE 1 (bases 1 to 134).  
 AUTHORS Morimyo,M. and Mita,K.  
 TITLE Identification of expressed sequence tags of Schizosaccharomyces pombe  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Mitsuoki Morimyo  
 Genome Research Group  
 National Institute of Radiological Sciences  
 9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan  
 Email: morimyo@nirs.go.jp.  
 Location/Qualifiers  
 1. .134  
 /organism="Schizosaccharomyces pombe"  
 /mol\_type="mRNA"  
 /strain="972"  
 /db\_xref="taxon:4896"  
 /clone="spc04760"  
 /sex="h minus"  
 /note="Vector: M13mp19; The cDNA library of Schizosaccharomyces pombe was prepared by cloning cDNA into the SmaI site of M13mp19 DNA and the direction of DNA sequences was not always from 5' to 3'. The cDNA data of Schizosaccharomyces pombe are available for searching on the World Wide Web. (URL, http://www.nirs.go.jp)"  
 US-10-030-194A-6 (1-6) x AU009304 (1-134)  
 QY 1 GlyTyr\*\*\*ValGlu 5  
 ||||| |||||  
 Db 16 GGATATTTCAGTAGAA 30  
 RESULT 42  
 EQ293027  
 LOCUS AU009307 136 bp mRNA linear EST 15-MAY-2002  
 DEFINITION MR1-AN0037-280800-003-d04 AN0037 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BQ293027  
 VERSION BQ293027.1 GI:20801977  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 136)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&t2=MR1-AN0037-280800-003-d04&t3=2000-08-28&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 16  
 High quality sequence stop: 136.  
 Location/Qualifiers  
 1. .136

9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan  
 Email: morimyo@nirs.go.jp.  
 Location/Qualifiers  
 1. .134  
 /organism="Schizosaccharomyces pombe"  
 /mol\_type="mRNA"  
 /strain="972"  
 /db\_xref="taxon:4896"  
 /clone="spc04760"  
 /sex="h minus"  
 /note="Vector: M13mp19; The cDNA library of Schizosaccharomyces pombe was prepared by cloning cDNA into the SmaI site of M13mp19 DNA and the direction of DNA sequences was not always from 5' to 3'. The cDNA data of Schizosaccharomyces pombe are available for searching on the World Wide Web. (URL, http://www.nirs.go.jp)"  
 US-10-030-194A-6 (1-6) x AU009304 (1-134)  
 QY 1 GlyTyr\*\*\*ValGlu 5  
 ||||| |||||  
 Db 16 GGATATTTCAGTAGAA 30  
 RESULT 42  
 EQ293027  
 LOCUS AU009307 136 bp mRNA linear EST 15-MAY-2002  
 DEFINITION MR1-AN0037-280800-003-d04 AN0037 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BQ293027  
 VERSION BQ293027.1 GI:20801977  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 136)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&t2=MR1-AN0037-280800-003-d04&t3=2000-08-28&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 16  
 High quality sequence stop: 136.  
 Location/Qualifiers  
 1. .136

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="AN0037"  
 /note="Organ: amion normal; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORSTES PCR (O.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.53e+04 Length: 136  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 5 Gaps: 0

US-10-030-194A-6 (1-6) x BQ293027 (1-136)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 52 GGGTATGCTGTCGAA 66

## RESULT 43

AI053196/C  
 LOCUS  
 DEFINITION TENU1620 T. cruzi epimastigote normalized cDNA Library Trypanosoma cruzi cDNA clone 19d11 3', mRNA sequence.  
 AI053196  
 ACCESSION AI053196.1 GI:3321075  
 VERSION  
 KEYWORDS EST.  
 SOURCE Trypanosoma cruzi

## ORGANISM

Trypanosoma cruzi  
 Trypanosoma cruzi  
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
 Trypanosoma; Schizotrypanum.  
 1 (bases 1 to 138)

## REFERENCE

AUTHORS Porcel,B.M., Tran,A.-N., Tammi,M., Nyarady,Z., Rydaker,M.,  
 Urmenyi,T.P., Rondinelli,E., Pettersson,U., Andersson,B. and  
 Aslund,L.

TITLE Gene survey of the pathogenic protozoan Trypanosoma cruzi

JOURNAL Genome Res. 10 (8), 1103-1107 (2000)

MEDLINE 20414748

PUBMED 10958628

## COMMENT

Contact: Aslund L,  
 Department of Medical Genetics  
 Uppsala University  
 Biomedical Center, Box 589, S-751 23 Uppsala, Sweden  
 Tel: 46 18 471 45 85  
 Fax: 46 18 52 68 49  
 Email: lena.aslund@medgen.uu.se  
 Seq primer: M13 forward  
 High quality sequence stop: 139.

## FEATURES

source  
 1..138  
 /organism="Trypanosoma cruzi"  
 /mol\_type="mRNA"  
 /strain="Cl-Brenner"  
 /db\_xref="taxon:5693"  
 /clone="19d11"  
 /cell\_type="epimastigote"  
 /clone\_lib="T. cruzi epimastigote normalized cDNA Library"  
 /note="cDNA library constructed with oligo dt primed  
 epimastigote mRNA and cloned in pT7t318D phagemid with  
 modified polylinker (PHARMACIA)"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.56e+04 Length: 138

Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 1 Gaps: 0

US-10-030-194A-6 (1-6) x AI053196 (1-138)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 55 GGGTACTCAGTAGAA 41

## RESULT 44

BQ912701

LOCUS

DEFINITION BQ912701 138 bp mRNA linear EST 19-AUG-2002  
 clone QHA22N01, mRNA sequence.

ACCESSION BQ912701

VERSION BQ912701.1 GI:22311480

KEYWORDS EST.

SOURCE Helianthus annuus (common sunflower)

## ORGANISM

Helianthus annuus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;  
 Heliantheae; Helianthus.

## REFERENCE

AUTHORS

1 (bases 1 to 138)  
 Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riesberg,L.,  
 Lin,H., van Damme,M., Lavalle,D., Chevalier,P., Ziegler,J.,  
 Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,  
 Lai,Z., Church,S., Jackson,L. and Bradford,K.  
 Lettuce and Sunflower ESTs from the Compositae Genome Project  
 http://compgenomics.ucdavis.edu/  
 Unpublished (2002)

## JOURNAL

COMMENT

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Amundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]

belongs to contig QH\_CA\_Contig2725, see http://cgpdb.ucdavis.edu/  
 for details.

Plate: QHA22 row: N column: 01.

## FEATURES

source

1..138  
 Location/Qualifiers  
 /organism="Helianthus annuus"  
 /mol\_type="mRNA"  
 /cultivar="RHA801"  
 /db\_xref="taxon:4232"  
 /clone="QHA22N01"  
 /lab\_host="E.coli"  
 /clone\_lib="OH ABCDI sunflower RHA801"  
 /note="Vector: pBRCNDSFIAB: The library was constructed  
 from 11 different sources of RNA from a single genotype.  
 Separate cDNAs were generated using primers that  
 incorporated unique 5' and 3' tags to distinguish each  
 source of RNA. cDNAs were then pooled, size-fractionated,  
 directionally cloned into a custom medium-copy vector and  
 transformations made with four size classes to minimize  
 size bias. Details of each source of RNA and library  
 construction can be obtained at http://cgpdb.ucdavis.edu/  
 TAG\_SEQ=Not found"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.56e+04 Length: 138  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 5 Gaps: 0

```

US-10-030-194A-6 (1-6) x BQ912701 (1-138)

Qy      1 GlyTyr***ValGlu 5
Db      27 GGGTACAGTGGGAA 41

RESULT 45
BW020881/c
LOCUS   BW020881 138 bp mRNA linear EST 13-OCT-2002
DEFINITION
intestinalis cDNA clone rcibd069e17 3', mRNA sequence.
ACCESSION
VERSION  BW020881.1 GI:23936688
KEYWORDS
SOURCE   EST.
ORGANISM
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE
1 (bases 1 to 138)
Satoch.Y., Satake,M., Azumi,K., Nonaka,M., Shin-i.T., Kohara,Y. and
Satoch.N.
TITLE    Expressed genes in Ciona intestinalis (2002)
JOURNAL  Unpublished (2002)
COMMENT  Contact: Nori Satoch
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satochascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..138
Location/Qualifiers
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcibd069e17"
/tissue_type="blood cells"
/clone_lib="Nori Satoch unpublished cDNA library, blood
cells"

ORIGIN
Alignment Scores:
Pred. No.: 1..56e+04 Length: 138
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 5 Gaps: 0

US-10-030-194A-6 (1-6) x BW020881 (1-138)

Qy      1 GlyTyr***ValGlu 5
Db      138 GGTATTCCGTCGAG 124

RESULT 46
AW750578
LOCUS   AW750578 140 bp mRNA linear EST 28-APR-2000
DEFINITION
RC0-CN0026-290100-011-cl2 CN0026 Homo sapiens cDNA, mRNA sequence.
ACCESSION
VERSION  AW750578.1 GI:7665510
KEYWORDS
SOURCE   EST.
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 140)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.

```

```

TITLE    Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE  20202663
PUBMED   10737800
COMMENT  Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC0&t2=RC0-CN0026-
290100-011-cl2&t3=2000-01-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 140.
Location/Qualifiers
1..140
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CN0026"
Note="Organ: colon normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Alignment Scores:
Pred. No.: 1..58e+04 Length: 140
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x AW750578 (1-140)

Qy      1 GlyTyr***ValGlu 5
Db      78 GGCTACACAGTGGAA 92

RESULT 47
AW582529
LOCUS   AW582529 141 bp mRNA linear EST 16-MAR-2000
DEFINITION
RC2-ST0256-110100-013-h06 ST0256 Homo sapiens cDNA, mRNA sequence.
ACCESSION
VERSION  AW582529.1 GI:7257578
KEYWORDS
SOURCE   EST.
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 141)
HCGP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

```



This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC2&t2=RC2-ST0256-110100-013-h06&t3=2000-01-11&t4=1)

Seq primer: puc 18 forward  
 High quality sequence start: 12  
 High quality sequence stop: 141.

#### FEATURES

source  
 1..141

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="ST0256"  
 /note="Organ: Stomach; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

#### ORIGIN

Alignment Scores:  
 Pred. No.: 1.6e+04 Length: 141  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x AW582529 (1-141)

#### Qy

1 GlyTyr\*\*\*ValGlu 5  
 |||||

Db 42 GGTACTCCGTAGAA 56

#### RESULT 48

BG930246/c

#### LOCUS

231-148 f231 S. mansoni adult mini-library, linear EST 03-JUL-2002  
 Fietto/DeMarco/Verjovski-Almeida Schistosoma mansoni cDNA, mRNA  
 sequence.

#### ACCESSION

VERSION BG930246.1 GI:17155183

#### KEYWORDS

EST.

#### SOURCE

Schistosoma mansoni  
 Schistosoma mansoni

#### REFERENCE

1 (bases 1 to 141)  
 Fietto J.L.R., DeMarco R. and Verjovski-Almeida S.

#### AUTHORS

Use of degenerate primers and touchdown PCR for construction of  
 cDNA libraries

#### TITLE

Biotechniques 32 (6), 1404-1408 (2002)

#### JOURNAL

Medline

#### PUBMED

12074173

#### COMMENT

Contact: Verjovski-Almeida S  
 Departamento de Bioquímica, Instituto de Química  
 Universidade de São Paulo  
 Av. Lineu Prestes, 748, São Paulo, SP 05508-900, Brasil  
 Tel: 55-11-3091-2173  
 Fax: 55-11-3091-2186  
 Email: verjovski@usp.br

#### FEATURES

source  
 1..141

/organism="Schistosoma mansoni"  
 /mol\_type="mRNA"  
 /strain="BH"

/db\_xref="taxon:6183"

/dev\_stage="Adult"

/clone\_lib="f231 S. mansoni adult mini-library,  
 Fietto/DeMarco/Verjovski-Almeida"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 1.6e+04 Length: 141  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x BG930246 (1-141)

#### Qy

1 GlyTyr\*\*\*ValGlu 5  
 |||||

Db 62 GGTATACGTGGAA 48

#### RESULT 49

CF023361

#### LOCUS

CF023361 OBO3e12.xg QBQ Zea mays cDNA clone QBQ3e12, mRNA sequence. EST 17-JUL-2003

#### ACCESSION

VERSION CF023361.1 GI:32918549

#### KEYWORDS

EST.

#### SOURCE

Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 141)

Genoplatte.

AUTHORS Genoplatte, a major partnership french program in plant genomics

TITLE Unpublished (2003)

JOURNAL Contact: Genoplatte

COMMENT Genoplatte

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french  
 plant genomics programme 'Genoplatte' (http://www.genoplatte.com  
 and http://genoplatte-info.inbio.fr).

#### FEATURES

source

1..141  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="F2"  
 /db\_xref="taxon:4577"  
 /clone="QBQ3e12"  
 /tissue\_type="endosperm"  
 /clone\_lib="QBQ"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 1.6e+04 Length: 141  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CF023361 (1-141)

#### Qy

1 GlyTyr\*\*\*ValGlu 5  
 |||||

Db 18 GGTACCGGTAGAA 32

#### RESULT 50

AW380229

#### LOCUS

AW380229 142 bp mRNA linear EST 04-PEB-2000

#### DEFINITION

QW3-HT0262-251099-007-c11 HT0262 Homo sapiens cDNA, mRNA sequence.

#### ACCESSION

AW380229

```

VERSION      AW380229.1  GI:6884797
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 142)
AUTHORS      HCGP http://www.ludwig.org.br/ORESTES.
TITLE        The FAPESP/LICR Human Cancer Genome Project
JOURNAL      Unpublished (1999)
COMMENT      Contact: Simpson A.J.G.
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922
             Fax: +55-11-2707001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the FAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&tl2=QV3-HT0262-251099-007-cll&t3=1999-10-25&t4=1)
             Seq primer: puc 18 forward
             High quality sequence start: 11
             High quality sequence stop: 107.

FEATURES             source
     source           1..142
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0262"
                     /note="Organ: head neck; Vector: puc18; Site 1: SmaI;
                     Site 2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the puc 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."

ORIGIN
Alignment Scores:      1.61e+04      Length:      142
Pred. No.:            22.00          Matches:      4
Score:                80.00%         Conservative: 0
Percent Similarity:   80.00%         Mismatches:   1
Best Local Similarity: 80.00%         Indels:       0
Query Match:         91.67%          Gaps:         0
DB:                  2

US-10-030-194A-6 (1-6) x AW380229 (1-142)

QY      1  GlyTyr***ValGlu 5
        |||||
        18 GGATACCTCAGTAGAA 32

Db

RESULT 51
CG602574
LOCUS     CG602574      142 bp      DNA      linear      GSS 02-OCT-2003
DEFINITION OST276104 Mus musculus 129Sv/Ev Mus musculus genomic clone
ACCESSION OST276104, genomic survey sequence.
VERSION    CG602574.1  GI:37422399
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 142)
Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
Payne,R.R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,


```

```

Spark,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
Zhu,Q., Person,C. and Sands,A.T.
Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
Contact: Zambrowicz BP
OmiBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.
Location/Qualifiers
     source           1..142
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129Sv/Ev"
                     /db_xref="taxon:10090"
                     /clone="OST276104"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129Sv/Ev"

ORIGIN
Alignment Scores:      1.61e+04      Length:      142
Pred. No.:            22.00          Matches:      4
Score:                80.00%         Conservative: 0
Percent Similarity:   80.00%         Mismatches:   1
Best Local Similarity: 80.00%         Indels:       0
Query Match:         91.67%          Gaps:         0
DB:                  0

US-10-030-194A-6 (1-6) x CG602574 (1-142)

QY      1  GlyTyr***ValGlu 5
        |||||
        72 GGGTATAGTGTGGAG 86

Db

RESULT 52
BE161529
LOCUS     BE161529      143 bp      mRNA      linear      EST 21-JUN-2000
DEFINITION RC3-HT0443-180200-012-e04 HT0443 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE161529
VERSION    BE161529.1  GI:8624250
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 143)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
PUBMED     10737800
COMMENT    Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922
             Fax: +55-11-2707001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the FAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC3-HT0443-180200-012-e04&t3=2000-02-18&t4=1)


```

```

Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 143.
Location/Qualifiers
1. .143
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0443"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Alignment Scores:
Pred. No.: 1.62e+04 Length: 143
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x BE161529 (1-143)

QY 1 GlyTyr***ValGlu 5
Db 22 GGGTATGCTGCGAA 36

RESULT 53
X83334/c 143 bp mRNA linear EST 27-SEP-1995
LOCUS MMSR42UNI Mouse cDNA Mus musculus cDNA clone RS42unito, mRNA
DEFINITION X83334
ACCESSION X83334.1 GI:949937
VERSION X83334.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Bettenhausen, B. and Gossler, A.
Efficient isolation of novel mouse genes differentially expressed
in early postimplantation embryos
JOURNAL Genomics 28 (3), 436-441 (1995)
MEDLINE 96039254
PUBMED 7490078
COMMENT Contact: B. Bettenhausen
Max-Deibueck-Laboratorium in der Max-Planck Gesellschaft
Carl-von-Linne-Weg 10, RL 50629 Koeln, FRG.
Location/Qualifiers
1. .143
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c x C57BL/6"
/db_xref="taxon:10090"
/clone_lib="Mouse cDNA"
/note="Vector: lambda ZAP II; Developmental stage = d 10.5
P.c."

ORIGIN
Alignment Scores:
Pred. No.: 1.62e+04 Length: 143
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0

Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 143.
Location/Qualifiers
1. .143
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0443"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Alignment Scores:
Pred. No.: 1.63e+04 Length: 144
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x BF843578 (1-144)

QY 1 GlyTyr***ValGlu 5

```

```
Db          90 CGTACTCTGTCGAA 76
|||||
|||||
|||||
|||||
|||||

RESULT 55
AW300193/c      144 bp      mRNA      linear      EST 18-JAN-2000
LOCUS          x660e05.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2774048 3'
DEFINITION     similar to TR:Q64157 Q64154 GENE TRAP PAT 12 ;, mRNA sequence.
ACCESSION      AW300193
VERSION        AW300193.1 GI:6709959
KEYWORDS
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 144)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL        Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
              Email: c9apbs-remail.nih.gov
              Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
              Emmert-Buck, M.D., Ph.D.
              cDNA Library Preparation: M. Bento Soares, Ph.D.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
FEATURES       source
              1..144
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:2774048"
                /lab_host="DH10B"
                /clone_lib="NCI CGAP Kid11"
                /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
                a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                Plasmid DNA from the normalized library NCI CGAP Kid3 was
                prepared, and as circles were made in vitro. Following HAP
                purification, this DNA was used as tracer in a subtractive
                hybridization reaction. The driver was PCR-amplified cDNAs
                from a pool of 5,000 clones made from the same library
                (clones 1322376-1323911, 1456007-1456775, and
                1500552-1502855). Subtraction by Bento Soares and M.
                Fatima Bonaldo."
ORIGIN
Alignment Scores:
Pred. No.:      1.63e+04      Length:      144
Score:          22.00         Matches:      4
Percent Similarity: 80.00%      Conservative: 0
Best Local Similarity: 80.00%      Mismatches:  1
Query Match:    91.67%         Indels:      0
DB:             2             Gaps:        0

US-10-030-194A-6 (1-6) x AW300193 (1-144)
QY      1 GlyTyr***ValGlu 5
|||||
Db      59 GGTATGTCAGTTGAG 45

RESULT 56
AW577008
LOCUS          PM3-BT0346-090100-001-e03 BT0346 Homo sapiens cDNA, mRNA sequence.
DEFINITION     PM3-BT0346-090100-001-e03 BT0346 Homo sapiens cDNA, mRNA sequence.
ACCESSION      AW577008
VERSION        AW577008
KEYWORDS
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 144)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL        Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
              Email: c9apbs-remail.nih.gov
              Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
              Emmert-Buck, M.D., Ph.D.
              cDNA Library Preparation: M. Bento Soares, Ph.D.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
FEATURES       source
              1..144
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:2774048"
                /lab_host="DH10B"
                /clone_lib="NCI CGAP Kid11"
                /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
                a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                Plasmid DNA from the normalized library NCI CGAP Kid3 was
                prepared, and as circles were made in vitro. Following HAP
                purification, this DNA was used as tracer in a subtractive
                hybridization reaction. The driver was PCR-amplified cDNAs
                from a pool of 5,000 clones made from the same library
                (clones 1322376-1323911, 1456007-1456775, and
                1500552-1502855). Subtraction by Bento Soares and M.
                Fatima Bonaldo."
ORIGIN
Alignment Scores:
Pred. No.:      1.63e+04      Length:      144
Score:          22.00         Matches:      4
Percent Similarity: 80.00%      Conservative: 0
Best Local Similarity: 80.00%      Mismatches:  1
Query Match:    91.67%         Indels:      0
DB:             2             Gaps:        0

US-10-030-194A-6 (1-6) x AW300193 (1-144)
QY      1 GlyTyr***ValGlu 5
|||||
Db      59 GGTATGTCAGTTGAG 45

RESULT 57
BH438734
LOCUS          BOHOS14TF BOHO Brassica oleracea genomic clone BOHOS14, genomic
DEFINITION     survey sequence.
ACCESSION      BH438734
VERSION        BH438734.1 GI:17624448
KEYWORDS       GSS.
SOURCE         Brassica oleracea
ORGANISM       Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 144)
AUTHORS        Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE          Whole genome shotgun sequencing of Brassica oleracea
JOURNAL        Unpublished (2001)
COMMENT        Other_GSSs: BOHOS14TR
```

```
VERSION      AW577008.1 GI:7252044
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 144)
AUTHORS        HCGP http://www.ludwig.org.br/ORESTES.
TITLE          The FAPESP/LICR Human Cancer Genome Project
JOURNAL        Unpublished (1999)
COMMENT        Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL
              (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PM3&t2=PM3-BT0346-
              090100-001-e03&t3=2000-01-09&t4=1)
              Seq primer: puc 18 forward
              High quality sequence start: 18
              High quality sequence stop: 144.
              Location/Qualifiers
FEATURES       source
              1..144
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /dev_stage="Adult"
                /clone_lib="BT0346"
                /note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:
                SmaI; A mini-library was made by cloning products derived
                from ORESTES PCR (U.S. Letters Patent application No.
                196,716 - Ludwig Institute for Cancer Research) profiles
                into the pUC 18 vector. Reverse transcription of tissue
                mRNA and cDNA amplification were performed under low
                stringency conditions."
ORIGIN
Alignment Scores:
Pred. No.:      1.63e+04      Length:      144
Score:          22.00         Matches:      4
Percent Similarity: 80.00%      Conservative: 0
Best Local Similarity: 80.00%      Mismatches:  1
Query Match:    91.67%         Indels:      0
DB:             2             Gaps:        0

US-10-030-194A-6 (1-6) x AW577008 (1-144)
QY      1 GlyTyr***ValGlu 5
|||||
Db      22 GGCTACACGGTGGAG 36

RESULT 57
BH438734
LOCUS          BOHOS14TF BOHO Brassica oleracea genomic clone BOHOS14, genomic
DEFINITION     survey sequence.
ACCESSION      BH438734
VERSION        BH438734.1 GI:17624448
KEYWORDS       GSS.
SOURCE         Brassica oleracea
ORGANISM       Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 144)
AUTHORS        Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE          Whole genome shotgun sequencing of Brassica oleracea
JOURNAL        Unpublished (2001)
COMMENT        Other_GSSs: BOHOS14TR
```

Contact: Chris Town

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF

Class: sheared ends.

Location/Qualifiers

#### FEATURES

1..144  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone\_lib="BOHOS14"  
/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"

#### ORIGIN

Alignment Scores:  
Pred. No.: 1.63e+04 Length: 144  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x BH438734 (1-144)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 86 GGGTACGCGGTGGAG 100

#### RESULT 58

BF368176

LOCUS CM3-GN0049-080900-338-c06\_1 GN0049 Homo sapiens cDNA, mRNA EST 24-NOV-2000  
DEFINITION sequence.

#### ACCESSION

BF368176

VERSION BF368176.1 GI:11330201

#### KEYWORDS

EST.

#### SOURCE

Homo sapiens (human)

#### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

#### REFERENCE

1 (bases 1 to 146)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM3&t2=CM3-GN0049-

080900-338-c06\_1&t3=2000-09-08&t4=1)

Seq primer: puc 18 forward.

Location/Qualifiers

#### FEATURES

source

1..146  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="GN0049"

/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

#### ORIGIN

Alignment Scores:  
Pred. No.: 1.66e+04 Length: 146  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x BF368176 (1-146)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 43 GGTATTTCAGTTGAG 57

#### RESULT 59

AV771220/c

LOCUS AV771220

DEFINITION AV771220 Lotus japonicus Pods (20-30 mm in length) Lotus corniculatus var. japonicus cDNA clone MPD018b08\_f 3', mRNA sequence.

#### ACCESSION

AV771220

VERSION AV771220.1 GI:45354467

#### KEYWORDS

EST.

#### SOURCE

Lotus corniculatus var. japonicus (Lotus japonicus)

Lotus corniculatus var. japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;

Lotus.

#### REFERENCE

1 (bases 1 to 148)

Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.

Characteristics of the Lotus japonicus Gene Repertoire Deduced from

Large-Scale Expressed Sequence Tag (EST) Analysis

Plant Mol. Biol. (2004) in press

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

#### FEATURES

1..148

Location/Qualifiers

source

/organism="Lotus corniculatus var. japonicus"

/mol\_type="mRNA"

/isolate="Miyakojima MG-20"

/db\_xref="taxon:34305"

/clone="MPD018b08\_f"

/tissue\_type="Pods (20-30 mm in length)"

/clone\_lib="Lotus japonicus Pods (20-30 mm in length)"

/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2: XhoI"

#### ORIGIN

Alignment Scores:

Pred. No.: 1.69e+04 Length: 148

Score: 22.00 Matches: 4

Percent Similarity: 80.00% Conservative: 0

Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0

DB: 1 Gaps: 0

US-10-030-194A-6 (1-6) x AV771220 (1-148)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 100 GGGTAGCGCGTGGAA 86

RESULT 60  
BE654176 148 bp mRNA linear EST 06-SEP-2000  
LOCUS  
DEFINITION  
UI-M-AN1-afc-g-04-0-UI.r1 NIH BMAP MBG N Mus musculus cDNA clone  
UI-M-AN1-afc-g-04-0-UI 5', mRNA sequence.

ACCESSION  
BE654176  
VERSION  
BE654176.1 GI:9980089

KEYWORDS  
EST.

SOURCE  
Mus musculus (house mouse)

ORGANISM  
Mus musculus

REFERENCE  
1 (bases 1 to 148)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)

JOURNAL  
77044477

MEDLINE  
889548

COMMENT  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mEST@mail.nih.gov  
cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It  
should be noted that Bento Soares is generating a small number of  
additional specialized non-redundant arrays of BMAP cDNAs whose  
availability will be considered under appropriate and limited  
collaborative arrangements  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1. .148  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-AN1-afc-g-04-0-UI"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NIH BMAP MBG N"  
/note="Vector: pT7T3D-PaC (Pharmacia) with a modified  
polylinker; Site 1: Not 1; Site 2: Eco RI; The  
NIH BMAP MBG N library is a normalized library constructed  
from mouse basal ganglia. The tag is a string of 5  
nucleotides present between the Not I site and the  
oligo-dT track. The library was constructed as described  
by Bonaldo, Lennon and Soares, Genome Research 6:  
791-806, 1996. Tissue provided by Ms. Annie Novakovich,  
Zivic-Miller Laboratories."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.69e+04 Length: 148  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
Gaps: 0

US-10-030-194A-6 (1-6) x BE654176 (1-148)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 100 GGCTACAGTGTGGAA 114

RESULT 61  
AZ461820 148 bp DNA linear GSS 04-OCT-2000  
LOCUS  
DEFINITION  
1M0267A24R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0267A24 R, genomic survey sequence.

ACCESSION  
AZ461820  
VERSION  
AZ461820.1 GI:10619945

KEYWORDS  
GSS.

SOURCE  
Mus musculus (house mouse)

ORGANISM  
Mus musculus

REFERENCE  
1 (bases 1 to 148)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
Reilly,M., Rose,M., Rose,R., Stokes,R., finge,A., von  
Niederhausern,A. and Wright,D.,Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)

JOURNAL  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: dunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0267 row: A column: 24  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 148.  
Location/Qualifiers  
1. .148  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0267A24"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptor DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptor mouse DNA was annealed to  
adaptor vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.69e+04 Length: 148  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0

```

DB:
US-10-030-194A-6 (1-6) x AZ461820 (1-148)
Gaps: 8 0

Qy 1 GlyTyr***ValGlu 5
|||||
69 GGTATACTGTGGAG 83

Db

RESULT 62
AZ749976
LOCUS
DEFINITION
  RPCI-24-78D12 TJ RPCI-24 Mus musculus genomic clone RPCI-24-78D12,
  genomic survey sequence.
ACCESSION
  AZ749976
VERSION
  AZ749976.1 GI:12535135
KEYWORDS
  GSS.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 148)
AUTHORS
  Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M.,
  Tsengave, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E.,
  Russell, D., de Jong, P. and Fraser, C.M.
  Mouse BAC End Sequences from Library RPCI-24
  Unpublished (1999)
  Contact: Shaying Zhao
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: szhao@tigr.org
  Clones are derived from the mouse BAC library RPCI-24. For BAC
  library availability, please contact Pieter de Jong
  (pdejong@mail.cho.org). Clones may be purchased from BACPAC
  Resources (http://www.choi.org/bacpac/orderingframe.html). BAC end
  page: http://www.tigr.org/tldb/bac\_ends/mouse/bac\_end\_intro.html
  Plate: 78 row: D column: 12
  Seq primer: SP6
  Class: BAC ends.
  Location/Qualifiers
    1..148
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /strain="C57BL/6J"
    /db_xref="taxon:10090"
    /clone="RPCI-24-78D12"
    /sex="Male"
    /cell_type="Spleen/Brain"
    /clone_lib="RPCI-24"
    /note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1;
    RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
    library was cloned in the pTARBAC1 cloning vector at the
    BamH1 sites using MboI partially digested male C57BL/6J
    DNA."

ORIGIN
Alignment Scores:
Pred. No.: 1.69e+04 Length: 148
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservativeness: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x AZ749976 (1-148)

Qy 1 GlyTyr***ValGlu 5
|||||
10 GGGTACTGTGGAG 24

Db

RESULT 63

```

```

CB817900
LOCUS
DEFINITION
  a6d02pz.r1 Plumbago zeylanica sperm cell (svm) cDNA library
  Plumbago zeylanica cDNA clone a6d02pz 5', mRNA sequence.
ACCESSION
  CB817900
VERSION
  CB817900.1 GI:29949876
KEYWORDS
  EST.
SOURCE
  Plumbago zeylanica
  Plumbago zeylanica
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Caryophyllales; Plumbaginaceae; Plumbago.
  1 (bases 1 to 150)
AUTHORS
  Gou, X.P., Yuan, T. and Russell, S.D.
  mRNAs of Plumbago zeylanica sperm cell populations
  Unpublished (2003)
  Contact: Russell, Scott
  Department of Botany and Microbiology
  University of Oklahoma
  Norman, OK 73019, USA
  Tel: 405 325 6234
  Fax: 405 325 7619
  Email: srussell@ou.edu
  Contact Dr. Scott D. Russell regarding clone availability
  Seq primer: PTRIP, 5'TCCGAGATCTGGACGAGC3'
  High quality sequence stop: 98.
  Location/Qualifiers
    1..150
    /organism="Plumbago zeylanica"
    /mol_type="mRNA"
    /db_xref="taxon:76149"
    /clone="a6d02pz"
    /cell_type="sperm cell"
    /dev_stage="post-anthesis pollen"
    /lab_host="E.coli strain XL1-Blue"
    /clone_lib="Plumbago zeylanica sperm cell (svm) cDNA
    library"
    /note="Vector: Clontech Triplex2; Sua and Svn sperm cells
    were isolated and collected in separate pools using a
    microinjector. cDNA libraries were constructed by oligo dT
    priming using the Clontech SMART cDNA Library Construction
    kit. The cDNAs were directionally cloned into vector
    Triplex2."

ORIGIN
Alignment Scores:
Pred. No.: 1.71e+04 Length: 150
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservativeness: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 6 Gaps: 0

US-10-030-194A-6 (1-6) x CB817900 (1-150)

Qy 1 GlyTyr***ValGlu 5
|||||
15 GGCTACAGTGTGGAA 29

Db

RESULT 64
AI459749/c
LOCUS
DEFINITION
  ap16h12.x1 Schiller oligodendrogliona Homo sapiens cDNA clone
  IMAGE:1955591 3', similar to SW:COX1 HUMAN P00395 CYTOCHROME C
  OXIDASE POLYPEPTIDE 1 ; mRNA sequence.
ACCESSION
  AI459749
VERSION
  AI459749.1 GI:4312630
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 151)
REFERENCE

```

**AUTHORS** Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, K. and Wilson, R. WashU-NCI human EST Project

**JOURNAL COMMENT** Unpublished (1997)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Trace considered overall poor quality  
Possible reversed clone: similarity on wrong strand  
Seq primer: -40UP from Gibco  
High quality sequence stop: 1.

**FEATURES** source  
1. .151  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1955591"  
/sex="male"  
/tissue type="oligodendrogloma"  
/dev\_stage="44 years"  
/lab\_host="SOUR"  
/clone\_lib="Schiller oligodendrogloma"  
/note="Organ: brain; Vector: pBluescript SK- (Stratagene); Site 1: EcoRI; Site 2: XhoI; Double-stranded cDNA was prepared from human oligodendrogloma using primer 5'-GAGAGAGAGAGAGAGAACTGCTGAGT(18)-3'. An EcoRI adaptor was used on the 5' end of the cDNA as follows: 5'-AATTCGGACAG-3'. The library was size-selected and went through one round of amplification. Average insert size is 1.7 kb, with a range from 0.4-12 kb. Tumor identification by consensus pathology; contains chromosome 1p and 19q deletion as determined by CGH. This library was constructed by Dr. Martin Schiller (Johns Hopkins University)."

**ORIGIN**

Alignment Scores:  
Pred. No.: 1.73e+04 Length: 151  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 1 Gaps: 0

US-10-030-194A-6 (1-6) x AI459749 (1-151)

**Qy** 1 GlyTyr\*\*\*ValGlu 5  
||||| |||||

**Db** 94 GGATACACCGTAGAG 80

**RESULT 65**  
BF047816  
LOCUS  
DEFINITION BF047816 151 bp mRNA linear EST 11-OCT-2000  
similar to TR:Q9ULX0 Q9ULX0 PYROPHOSPHATASE. ;, mRNA sequence.

**ACCESSION** BF047816  
**VERSION** BF047816.1 GI:10766319  
**KEYWORDS** EST.  
**SOURCE** Xenopus laevis (African clawed frog)  
**ORGANISM** Xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Xenopus.

**REFERENCE** 1 (bases 1 to 151)  
**AUTHORS** NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

**JOURNAL COMMENT** Tumor Gene Index  
Unpublished (1997)  
Other ESTs: dc83e04.x1  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov  
Trace considered overall poor quality  
Seq primer: -40RP from Gibco  
High quality sequence stop: 1.

**FEATURES** source  
1. .151  
Location/Qualifiers  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:3403710"  
/tissue type="oocytes"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NICHDXGC-001"  
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;  
Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.2 kb. Constructed by Life Technologies."

**ORIGIN**

Alignment Scores:  
Pred. No.: 1.73e+04 Length: 151  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x BF047816 (1-151)

**Qy** 1 GlyTyr\*\*\*ValGlu 5  
||||| |||||

**Db** 5 GGCTACACCGTAGAG 19

**RESULT 66**  
AV017443  
LOCUS  
DEFINITION AV017443 152 bp mRNA linear EST 28-AUG-1999  
clone 1110070M13, mRNA sequence.

**ACCESSION** AV017443  
**VERSION** AV017443.1 GI:4794435  
**KEYWORDS** EST.  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM** Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE** 1 (bases 1 to 152)  
**AUTHORS** Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K., Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Hori, F., Ighikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Niituma, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

**TITLE** RIKEN Mouse ESTs  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Chie Owa  
Genome Science Laboratory  
RIKEN  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-9145  
Fax: 81-298-36-9098  
Email: genome-res@tc.riken.go.jp  
Thermotabilization and thermoactivation of thermolabile enzymes by



trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))  
 Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

## FEATURES

source

Location/Qualifiers  
 1..152  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="1110070M13"  
 /sex="mixed"  
 /dev\_stage="18-day embryo"  
 /clone\_lib="Mus musculus 18-day embryo C57BL/6J"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.74e+04 Length: 152  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 1 Gaps: 0

US-10-030-194A-6 (1-6) x AV017443 (1-152)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 7 GGGATGCGAGTGGAG 21

## RESULT 67

N76484/c

LOCUS

DEFINITION N76484 152 bp mRNA linear EST 02-APR-1996  
 y287h10.r1 Soares multiple sclerosis 2NbHMSF Homo sapiens cDNA  
 clone IMAGE:290083 5', mRNA sequence.

ACCESSION N76484

VERSION N76484.1

KEYWORDS GI:1239062

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## AUTHORS

1 (bases 1 to 152)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, B., Rohlfing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project

JOURNAL

COMMENT

Unpublished (1995)  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

This clone is available royalty-free through LNL ; contact the IMAGE Consortium ([infoimage.lnl.gov](http://infoimage.lnl.gov)) for further information.

Seq primer: reverse ET

High quality sequence stop: 137.

Location/Qualifiers

source

1..152  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:3905923"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:290083"  
 /sex="male"  
 /tissue\_type="multiple sclerosis lesions"  
 /dev\_stage="Age 46"  
 /lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares multiple sclerosis 2NbHMSF"  
 /note="Vector: pT7T3D (Pharmacia) with a modified polylinker V-type: Phagemid; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.74e+04 Length: 152  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 7 Gaps: 0

US-10-030-194A-6 (1-6) x N76484 (1-152)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 50 GGATATGCTGTAGAA 36

## RESULT 68

BI045919

LOCUS

DEFINITION

MR3-FN0206-020201-008-g12

ACCESSION

BI045919

VERSION

BI045919.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 153)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsuoka, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3&t2=MR3-FN0206-020201-008-g12&t3=2001-02-02&t4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 152.

Location/Qualifiers

source

1..153

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

BI045919 153 bp mRNA linear EST 14-JUN-2001  
 MR3-FN0206-020201-008-g12 FN0206 Homo sapiens cDNA, mRNA sequence.

ACCESSION BI045919

VERSION BI045919.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 153)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsuoka, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3&t2=MR3-FN0206-020201-008-g12&t3=2001-02-02&t4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 152.

Location/Qualifiers

source

1..153

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="FN0206"  
 /note="Organ: prostate normal; Vector: puc18; Site 1:  
 Sma1; Site 2: Sma1; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent  
 application No. 196,716 - Ludwig Institute for Cancer  
 Research) profiles into the pUC 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.75e+04 Length: 153  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x BI045919 (1-153)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 5 GGATATAGCGGTGAG 19

## RESULT 69

LOCUS BQ255119 153 bp mRNA linear EST 06-MAY-2002  
 DEFINITION NISC\_Jp12h04.y1 NCI\_CGAP\_BCl Mus musculus cDNA clone IMAGE:4086775  
 5', mRNA sequence.

ACCESSION BQ255119  
 VERSION BQ255119.1 GI:20450932  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 153)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL Tumor Gene Index

COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

cDNA Library Preparation: David B. Krizman, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC)

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

[info@image.llnl.gov](mailto:info@image.llnl.gov)

Plate: L1AM9351 row: P column: 8

Seq primer: M13RPI reverse primer (ABI).

## FEATURES

source  
 1..153  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4086775"  
 /tissue\_type="flow-sorted, common myeloid progenitors"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP BCl"  
 /note="Organ: blood; Vector: pAMP1; mRNA made from myeloid  
 progenitors, cDNA made by oligo-dT priming. Directionally  
 cloned into UDG sites. Size-selected on agarose gel,  
 average insert size 300 bp. Primary library. cDNA Library  
 Preparation: David B. Krizman, Ph.D."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.75e+04 Length: 153  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0  
 DB: 5 Gaps: 0

US-10-030-194A-6 (1-6) x BQ255119 (1-153)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 82 GGCTACACAGTGGAG 96

## RESULT 70

LOCUS CR066309 153 bp DNA linear GSS 05-JUL-2004  
 DEFINITION Reverse strand read from insert in 3'HPRT insertion targeting and  
 chromosome engineering clone MHPPI92n10, genomic survey sequence.

ACCESSION CR066309

VERSION CR066309.1 GI:49799899

KEYWORDS GSS; genome survey sequence; MICER.

SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 153)

AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
 Rogers,J. and Bradley,A.

TITLE Direct Submission

JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>

FEATURES Location/Qualifiers

source  
 1..153  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="MHPPI92n10"  
 /clone\_lib="MHPPI"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.75e+04 Length: 153  
 Score: 22.00 Matches: 4  
 Percent Similarity: 80.00% Conservative: 0  
 Best Local Similarity: 80.00% Mismatches: 1  
 Query Match: 91.67% Indels: 0  
 DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x CR066309 (1-153)

Qy 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 Db 64 GGATATAGTGTGGAA 78

## RESULT 71

LOCUS AV950897/c

DEFINITION AV950897 Nori Satoh unpublished cDNA library, cleavage stage embryo  
 Ciona intestinalis cDNA clone c1c105f17 5', mRNA sequence.

ACCESSION AV950897

VERSION AV950897.1 GI:19439196

KEYWORDS EST.

SOURCE Ciona intestinalis

## ORGANISM

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
 Phlebobranchia; Cionidae; Ciona.

REFERENCE 1 (bases 1 to 155)

AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished (2000)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoheascidian.zool.kyoto-u.ac.jp.

# FEATURES

source

1. .155  
Location/Qualifiers  
/organism="Ciona intestinalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:7719"  
/clone="cicl05f17"  
/tissue\_type="whole animal"  
/dev\_stage="cleavage stage embryo"  
/clone\_lib="Nori Satoch unpublished cDNA library, cleavage stage embryo"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.78e+04 Length: 155  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 1 Gaps: 0

US-10-030-194A-6 (1-6) x AV950897 (1-155)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 109 GGTATGCTGTGAG 95

## RESULT 72

BF806750 155 bp mRNA linear EST 12-JAN-2001  
LOCUS RC2-CI0080-081100-016-e10 CI0080 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF806750

VERSION BF806750.1 GI:12135648

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 155)

AUTHORS  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC2&t2=RC2-CI0080-081100-016-e10&t3=2000-11-08&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 23

High quality sequence stop: 155.

Location/Qualifiers

1. .155

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="CI0080"

/note="Organ: colon\_ins; Vector: puc18; Site\_1: SmaI;

## FEATURES

source

1. .155  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT0342"  
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No.196,716 - Ludwig Institute for Cancer Research)  
No.196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under

## ORIGIN

Alignment Scores:  
Pred. No.: 1.78e+04 Length: 155  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x BF806750 (1-155)

Qy 1 GlyTyr\*\*\*ValGlu 5

Db 20 GGTACAGTGTGAA 34

## RESULT 73

BE154383/c

LOCUS PM2-HT0342-081299-001-h07 HT0342 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE154383

VERSION BE154383.1 GI:8617104

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 155)

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=8t2&t2=PM2-HT0342-081

299-001-h07&t3=1999-12-08&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 18

High quality sequence stop: 155.

Location/Qualifiers

1. .155

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT0342"

low stringency conditions."

ORIGIN

Alignment Scores: 1.78e+04 Length: 155  
Pred. No.: 22.00 Matches: 4  
Score: 22.00 Conservativeness: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 2

US-10-030-194A-6 (1-6) x BE154383 (1-155)

QY 1 GlyTyr\*\*\*ValGlu 5

DB 143 GGGTACTCCGTGGAG 129

RESULT 74

BM986915 156 bp mRNA linear EST 15-JUL-2003  
LOCUS UI-H-C00-aqj-f-11-0-UI.s1 NCI\_CGAP\_Sub9 Homo sapiens cDNA clone  
DEFINITION IMAGE:3104372 3', mRNA sequence.

ACCESSION BM986915  
VERSION BM986915.1 GI:19706304

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 156)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

RNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

Clone Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

The following repetitive elements were found in this cDNA

sequence: 47-127. >HY3#scrna

Seq primer: M13 FORWARD

POLYA=yes.

FEATURES

source

1. .156  
Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3104372"

/tissue\_type="mixed"

/dev\_stage="mixed"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NCI CGAP Sub9"

/note="Vector: pT7T3-Pac (Pharmacia) with a modified

polylinker; Site 1: EcoR I; Site 2: Not I; tissues:

Cholonic mucosa with Crohns disease, Cholonic mucosa with

ulcerative colitis, Fetal thymus, Cervix, Cervical

adenosquamous carcinoma, Ligament cells, Prostate

carcinoma, Bladder carcinoma, Brain oligodendroglia ;

NCI CGAP Sub9 is a substracted cDNA library constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dr primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT7T3-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tags for this library are CGTC, AACG, GGCC,

GGAAG, TAGC, ATGG, AGACA, ATCAC. For additional

information, contact: Bento Soares, bento-soares@uiowa.edu  
TAG\_TISSUE=Brain Oligodendroglia  
TAG\_LIB=UI-H-C00  
TAG\_SEQ=ATCAC"

ORIGIN

Alignment Scores: 1.79e+04 Length: 156  
Pred. No.: 22.00 Matches: 4  
Score: 22.00 Conservativeness: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 5

US-10-030-194A-6 (1-6) x BM986915 (1-156)

QY 1 GlyTyr\*\*\*ValGlu 5

DB 133 GGGTACGCGAGTGGAG 119

RESULT 75

CN945347 156 bp mRNA linear EST 07-JUN-2004  
LOCUS 011010AVBC224045PG (AVBC) Royal Gala young shoot Malus x domestica  
DEFINITION cDNA clone AVBC224045, mRNA sequence.

ACCESSION CN945347  
VERSION CN945347.1 GI:48418160

KEYWORDS EST.

SOURCE Malus x domestica (cultivated apple)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eustosids I; Rosales; Rosaceae; Maloideae; Malus.

REFERENCE 1 (bases 1 to 156)

AUTHORS

Beuning, L., Bowen, J., Crowhurst, R., Gleave, A., Janssen, B.,

McCartney, S., Newcomb, R., Ross, G., Snowden, K., Walton, E. and Yauk, Y.

HortResearch Apple EST Project

Unpublished (2004)

Contact: Gleave, A.

Sequencing Facility

The Horticulture and Food Research Institute of New Zealand Ltd

120 Mt Albert Rd, Mt Albert, Auckland, New Zealand

Tel: 00 64 09 815 4200

Fax: 00 64 09 815 4201

Email: [est@hortresearch.co.nz](mailto:est@hortresearch.co.nz).

FEATURES

source

1. .156  
Location/Qualifiers

/organism="Malus x domestica"

/mol\_type="mRNA"

/db\_xref="taxon:3750"

/clone="AVBC224045"

/tissue\_type="Vegetative bud"

/dev\_stage="Pre-opening"

/clone\_lib="(AVBC) Royal Gala young shoot"

/note="Vector: pBluescript SK(-); Library sequenced by 1.

Genesis Research & Development 2. Paradigm Genetics"

ORIGIN

Alignment Scores: 1.79e+04 Length: 156  
Pred. No.: 22.00 Matches: 4  
Score: 22.00 Conservativeness: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 7

US-10-030-194A-6 (1-6) x CN945347 (1-156)

QY 1 GlyTyr\*\*\*ValGlu 5

DB 55 GGGTACGCGGTGGAA 69

Search completed: November 3, 2004, 17:55:29  
Job time : 2576 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 3, 2004, 13:41:12 ; Search time 72.5 Seconds  
(without alignments)  
58.824 Million cell updates/sec

Title: US-10-030-194A-6

Perfect score: 24  
Sequence: 1 GYXVEX 6

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO pool\_p/US10030194/runat\_01112004.184848.28987/app\_query.fasta\_1.398  
-DB=Issued Patents NA -QFNT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=1000 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75  
-MODE=LOCAL -OUTFMT=pfo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10030194 @CGN 1.1.164 @runat 01112004.184848.28987 -NCPU=6 -ICPU=3  
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	91.7	16	2	US-08-447-173A-64
C 2	22	91.7	21	4	US-09-380-836-55
C 3	22	91.7	21	4	US-09-380-836-55
C 4	22	91.7	48	2	US-08-865-675-7
C 5	22	91.7	48	2	US-09-237-510-7
C 6	22	91.7	75	4	US-08-956-171E-3338
C 7	22	91.7	75	4	US-08-781-986A-3338
C 8	22	91.7	95	2	US-08-353-476-61
C 9	22	91.7	100	3	US-09-298-886-14
10	22	91.7	100	3	US-09-298-886-17
11	22	91.7	100	4	US-09-999-672-14
12	22	91.7	100	4	US-09-999-672-17

13	22	91.7	166	4	US-08-956-171E-4703	Sequence 4703, Ap
14	22	91.7	166	4	US-08-781-986A-4703	Sequence 4703, Ap
15	22	91.7	182	4	US-09-270-767-1906	Sequence 1906, Ap
16	22	91.7	182	4	US-09-270-767-17188	Sequence 17188, A
c 17	22	91.7	193	3	US-09-063-733A-53	Sequence 53, Appl
c 18	22	91.7	202	4	US-09-016-434-285	Sequence 285, App
c 19	22	91.7	206	4	US-09-513-999C-35499	Sequence 35499, A
c 20	22	91.7	225	4	US-09-134-000C-2517	Sequence 2517, Ap
21	22	91.7	228	4	US-09-134-000C-1406	Sequence 1406, Ap
22	22	91.7	273	4	US-09-696-569-1	Sequence 1, Appli
23	22	91.7	277	4	US-09-513-999C-28003	Sequence 28003, A
c 24	22	91.7	285	4	US-09-513-999C-26667	Sequence 26667, A
c 25	22	91.7	286	2	US-08-447-173A-59	Sequence 59, Appl
c 26	22	91.7	286	2	US-08-447-173A-60	Sequence 60, Appl
c 27	22	91.7	292	4	US-09-720-655B-3	Sequence 3, Appli
28	22	91.7	314	1	US-09-609-324A-7	Sequence 7, Appli
29	22	91.7	314	1	US-08-920-440B-7	Sequence 7, Appli
30	22	91.7	314	3	US-09-173-492-7	Sequence 7, Appli
31	22	91.7	314	3	US-08-173-133-7	Sequence 7, Appli
32	22	91.7	314	3	US-09-165-533-7	Sequence 7, Appli
33	22	91.7	314	3	US-09-580-236A-7	Sequence 7, Appli
34	22	91.7	322	3	US-09-323-873A-9	Sequence 9, Appli
35	22	91.7	326	4	US-08-956-171E-4450	Sequence 4450, Ap
36	22	91.7	326	4	US-08-781-986A-4450	Sequence 4450, Ap
c 37	22	91.7	342	3	US-08-929-856-65	Sequence 65, Appl
c 38	22	91.7	342	3	US-08-929-856-188	Sequence 188, App
39	22	91.7	342	4	US-09-134-000C-1407	Sequence 1407, Ap
c 40	22	91.7	348	4	US-09-252-991A-12495	Sequence 12495, A
c 41	22	91.7	348	4	US-09-134-000C-2798	Sequence 2798, Ap
c 42	22	91.7	366	3	US-09-030-607-215	Sequence 215, App
c 43	22	91.7	366	3	US-09-439-313-215	Sequence 215, App
c 44	22	91.7	366	3	US-09-352-616A-215	Sequence 215, App
c 45	22	91.7	366	4	US-09-232-149A-215	Sequence 215, App
c 46	22	91.7	366	4	US-09-159-812-215	Sequence 215, App
c 47	22	91.7	366	4	US-09-636-215-215	Sequence 215, App
c 48	22	91.7	366	4	US-09-688-166A-215	Sequence 215, App
c 49	22	91.7	366	4	US-09-115-453-215	Sequence 215, App
c 50	22	91.7	366	4	US-09-688-489-215	Sequence 215, App
c 51	22	91.7	366	4	US-09-679-426-215	Sequence 215, App
c 52	22	91.7	372	2	US-08-454-557C-35	Sequence 35, Appl
53	22	91.7	372	2	US-08-340-426D-35	Sequence 35, Appl
54	22	91.7	372	2	US-08-450-673C-35	Sequence 35, Appl
55	22	91.7	372	5	PCT-US95-1711A-35	Sequence 35, Appl
c 56	22	91.7	384	4	US-09-702-705-1768	Sequence 1768, Ap
c 57	22	91.7	384	4	US-09-736-457-1768	Sequence 1768, Ap
c 58	22	91.7	384	4	US-09-671-325-1768	Sequence 1768, Ap
c 59	22	91.7	384	4	US-09-658-824-1768	Sequence 1768, Ap
c 60	22	91.7	405	2	US-08-621-751A-11	Sequence 11, Appl
61	22	91.7	414	4	US-09-489-039A-1457	Sequence 1457, Ap
c 62	22	91.7	418	4	US-09-976-594-512	Sequence 512, App
c 63	22	91.7	420	4	US-09-134-000C-749	Sequence 749, App
64	22	91.7	424	4	US-09-621-976-17465	Sequence 17465, A
c 65	22	91.7	427	4	US-09-621-976-10373	Sequence 10373, A
c 66	22	91.7	432	4	US-08-252-991A-16414	Sequence 16414, A
67	22	91.7	472	4	US-08-956-171E-3176	Sequence 3176, Ap
68	22	91.7	472	4	US-08-781-986A-3176	Sequence 3176, Ap
69	22	91.7	475	4	US-09-621-976-2058	Sequence 2058, Ap
c 70	22	91.7	484	4	US-09-621-976-14205	Sequence 14205, A
c 71	22	91.7	486	4	US-09-252-991A-8516	Sequence 8516, Ap
c 72	22	91.7	489	4	US-09-270-767-3772	Sequence 3772, Ap
c 73	22	91.7	489	4	US-09-270-767-19054	Sequence 19054, A
74	22	91.7	495	4	US-09-134-000C-1338	Sequence 1338, Ap
75	22	91.7	513	4	US-09-621-976-3143	Sequence 3143, Ap
76	22	91.7	515	4	US-09-621-976-16591	Sequence 16591, A
77	22	91.7	521	3	US-09-323-873A-7	Sequence 7, Appli
78	22	91.7	555	1	US-09-609-324A-1	Sequence 1, Appli
79	22	91.7	555	3	US-08-920-440B-1	Sequence 1, Appli
80	22	91.7	555	3	US-09-173-492-1	Sequence 1, Appli
81	22	91.7	555	3	US-09-173-133-1	Sequence 1, Appli
82	22	91.7	555	3	US-09-165-533-1	Sequence 1, Appli
83	22	91.7	555	4	US-09-580-236A-1	Sequence 1, Appli
84	22	91.7	555	5	PCT-US95-12779-1	Sequence 1, Appli
85	22	91.7	555	5	PCT-US95-15781-1	Sequence 1, Appli















962 21 87.5 966 3 US-09-361-434-19  
c 963 21 87.5 966 3 US-09-361-434-20  
964 21 87.5 966 3 US-09-635-025-19  
c 965 21 87.5 966 3 US-09-635-025-20  
966 21 87.5 972 1 US-08-404-445-2  
c 967 21 87.5 972 3 US-09-134-001C-296  
968 21 87.5 972 4 US-09-891-641-9  
969 21 87.5 975 4 US-09-328-352-510  
c 970 21 87.5 981 4 US-09-671-317-38  
971 21 87.5 981 4 US-09-614-221A-555  
c 972 21 87.5 984 4 US-09-248-796A-6727  
973 21 87.5 989 3 US-09-361-434-21  
c 974 21 87.5 989 3 US-09-361-434-23  
975 21 87.5 989 3 US-09-635-025-21  
c 976 21 87.5 989 3 US-09-635-025-23  
977 21 87.5 993 4 US-09-252-991A-13190  
978 21 87.5 993 4 US-09-107-532A-3251  
979 21 87.5 993 4 US-09-134-000C-86  
980 21 87.5 993 4 US-09-248-796A-6692  
981 21 87.5 997 4 US-09-143-476-307  
982 21 87.5 999 3 US-08-963-901-3  
983 21 87.5 999 4 US-09-792-024-11  
984 21 87.5 1001 4 US-09-641-638-204  
c 985 21 87.5 1001 4 US-09-671-317-401  
c 986 21 87.5 1001 4 US-09-671-317-402  
987 21 87.5 1001 4 US-09-671-317-442  
c 988 21 87.5 1001 4 US-09-671-317-446  
989 21 87.5 1001 4 US-10-170-097-204  
c 990 21 87.5 1002 4 US-09-252-991A-12765  
991 21 87.5 1002 4 US-09-543-681A-3396  
c 992 21 87.5 1011 3 US-08-975-762-18  
993 21 87.5 1011 3 US-08-821-324-18  
c 994 21 87.5 1011 3 US-09-295-028-18  
c 995 21 87.5 1011 3 US-09-106-582-18  
c 996 21 87.5 1011 4 US-09-159-469-18  
c 997 21 87.5 1011 4 US-09-693-542-18  
998 21 87.5 1011 4 US-09-270-767-10840  
999 21 87.5 1014 4 US-09-252-991A-8083  
1000 21 87.5 1020 4 US-09-328-352-548

ALIGNMENTS

RESULT 1  
US-08-447-173A-64/c  
; Sequence 64, Application US/08447173A  
; Patent No. 5843730  
; GENERAL INFORMATION:  
; APPLICANT: Wain-Hobson, Simon  
; APPLICANT: Pezo, Valerie  
; TITLE OF INVENTION: Method For Hypermutagenesis Of  
; TITLE OF INVENTION: Nucleotides  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I. Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,173A  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146

; REFERENCE/DOCKET NUMBER: 03495.0142-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)408-4000  
; TELEFAX: (202)408-4400  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-447-173A-64  
Alignment Scores: 26.2 Length: 16  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 91.67% Indels: 0  
Query Match: 2 Gaps: 0  
DB:  
US-10-030-194A-6 (1-6) x US-08-447-173A-64 (1-16)  
Oy 1 GlyTyr\*\*\*ValGlu 5  
Db 15 GGCTACGCGGTCGAG 1  
RESULT 2  
US-09-380-836-55/c  
; Sequence 55, Application US/09380836  
; Patent No. 6551775  
; GENERAL INFORMATION:  
; APPLICANT: Lifton, Richard P.  
; APPLICANT: Chang, Sue S.  
; APPLICANT: Rossier, Bernard C.  
; TITLE OF INVENTION: Method to Diagnose and Treat Pathological Conditions  
; TITLE OF INVENTION: Resulting from Deficient Ion Transport such as  
; FILE REFERENCE: 44574-5018-US  
; CURRENT APPLICATION NUMBER: US/09/380,836  
; CURRENT FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/040,171  
; PRIOR FILING DATE: 1997-03-11  
; PRIOR APPLICATION NUMBER: PCT/US98/04681  
; PRIOR FILING DATE: 1998-03-11  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 55  
; LENGTH: 21  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: B-3 forward  
; OTHER INFORMATION: PCR primer  
US-09-380-836-55  
Alignment Scores: 35.2 Length: 21  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 91.67% Indels: 0  
Query Match: 4 Gaps: 0  
DB:  
US-10-030-194A-6 (1-6) x US-09-380-836-55 (1-21)  
Oy 1 GlyTyr\*\*\*ValGlu 5  
Db 19 GGCTACACTGTGGAG 5  
RESULT 3  
US-09-380-836-56/c  
; Sequence 56, Application US/09380836

```
; Patent No. 6551775
; GENERAL INFORMATION:
; APPLICANT: Lifton, Richard P.
; APPLICANT: Chang, Sue S.
; APPLICANT: Rossier, Bernard C.
; TITLE OF INVENTION: Method to Diagnose and Treat Pathological Conditions
; TITLE OF INVENTION: Resulting from Deficient Ion Transport such as
; TITLE OF INVENTION: Pseudohypoadosteronism Type-1
; FILE REFERENCE: 44574-5018-US
; CURRENT APPLICATION NUMBER: US/09/380,836
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/040,171
; PRIOR FILING DATE: 1997-03-11
; PRIOR APPLICATION NUMBER: PCT/US98/04681
; PRIOR FILING DATE: 1998-03-11
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: B-4 forward
; OTHER INFORMATION: PCR primer
US-09-380-836-56

Alignment Scores:
Pred. No.: 35.2 Length: 21
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-380-836-56 (1-21)

Qy 1 GlyTyr***ValGlu 5
Db 19 GGCTACACTGTGGAG 5

RESULT 4
US-08-865-675-7/c
; Sequence 7, Application US/08865675
; Patent No. 5928869
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Pitner, James B.
; APPLICANT: Linn, Carl P.
; APPLICANT: Schram, James L.
; TITLE OF INVENTION: DETECTION OF NUCLEIC ACIDS BY
; TITLE OF INVENTION: FLUORESCENCE QUENCHING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R. J. Rodrick, Becton Dickinson and Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,675
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3746
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-865-675-7

Alignment Scores:
Pred. No.: 86.6 Length: 48
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-237-510-7 (1-48)

Qy 1 GlyTyr***ValGlu 5
Db 44 GGTACTCAGTAGAG 30

RESULT 5
US-09-237-510-7/c
; Sequence 7, Application US/09237510
; Patent No. 5958700
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Pitner, James B.
; APPLICANT: Linn, Carl P.
; APPLICANT: Schram, James L.
; TITLE OF INVENTION: DETECTION OF NUCLEIC ACIDS BY
; TITLE OF INVENTION: FLUORESCENCE QUENCHING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R. J. Rodrick, Becton Dickinson and Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/237,510
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3746
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-237-510-7

Alignment Scores:
Pred. No.: 86.6 Length: 48
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-237-510-7 (1-48)

Qy 1 GlyTyr***ValGlu 5
Db 44 GGTACTCAGTAGAG 30
```





; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bencen, Gerard H  
; REGISTRATION NUMBER: 35,746  
; REFERENCE/DOCKET NUMBER: GP-100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (904) 375-8100  
; TELEFAX: (904) 372-5800  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 95 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-353-476-61

Alignment Scores:  
Pred. No.: 182 Length: 95  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x US-08-353-476-61 (1-95)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 20 GGATATACAGTGGAA 34

## RESULT 9

US-09-298-886-14  
; Sequence 14, Application US/09298886  
; Patent No. 6329170  
; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; FILE REFERENCE: THEREOF

; CURRENT APPLICATION NUMBER: US/09/298,886  
; PRIOR FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14

; LENGTH: 100  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-298-886-14

Alignment Scores:  
Pred. No.: 193 Length: 100  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-298-886-14 (1-100)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 11 GGCTACACCGTGGAA 25

## RESULT 10

US-09-298-886-17  
; Sequence 17, Application US/09298886  
; Patent No. 6329170  
; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; FILE REFERENCE: THEREOF  
; FILE REFERENCE: 8511-029  
; CURRENT APPLICATION NUMBER: US/09/298,886  
; CURRENT FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 100  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-298-886-17

Alignment Scores:  
Pred. No.: 193 Length: 100  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-298-886-17 (1-100)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 11 GGCTACACCGTGGAA 25

## RESULT 11

US-09-999-672-14  
; Sequence 14, Application US/09999672  
; Patent No. 6656714  
; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; FILE REFERENCE: THEREOF

; CURRENT APPLICATION NUMBER: US/09/999,672  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US/09/298,886  
; PRIOR FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14

; LENGTH: 100  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-999-672-14

Alignment Scores:  
Pred. No.: 193 Length: 100  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-999-672-14 (1-100)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 11 GGCTACACCGTGGAA 25

## RESULT 12

US-09-999-672-17  
; Sequence 17, Application US/09999672  
; Patent No. 6656714  
; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

```

; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-999-672-17

```

```

Alignment Scores:
Pred. No.: 193 Length: 100
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

```

US-10-030-194A-6 (1-6) x US-09-999-672-17 (1-100)

```

QY 1 GlyTyr***ValGlu 5
DB 11 GGCTACACCGTGGAA 25

```

```

RESULT 13
US-08-956-171E-4703
; Sequence 4703, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 4703:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4703:
US-08-956-171E-4703

```

```

Alignment Scores:
Pred. No.: 335 Length: 166
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

```

US-10-030-194A-6 (1-6) x US-08-956-171E-4703 (1-166)

```

QY 1 GlyTyr***ValGlu 5
DB 13 GGTATTCGTGAA 27

```

```

RESULT 14
US-08-781-986A-4703
; Sequence 4703, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 4703:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-4703

```

```

Alignment Scores:
Pred. No.: 335 Length: 166
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

```

US-10-030-194A-6 (1-6) x US-08-781-986A-4703 (1-166)

```

QY 1 GlyTyr***ValGlu 5

```

```
Db 13 GGTATTCTGTGAA 27

RESULT 15
US-09-270-767-1906
; Sequence 1906, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1906
; LENGTH: 182
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1906

Alignment Scores:
Pred. No.: 370 Length: 182
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-270-767-1906 (1-182)

QY 1 GlyTyr***ValGlu 5
|||||
DB 86 GGCTACTCGGTGAG 100

RESULT 16
US-09-270-767-17188
; Sequence 17188, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17188
; LENGTH: 182
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17188

Alignment Scores:
Pred. No.: 370 Length: 182
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-270-767-17188 (1-182)

QY 1 GlyTyr***ValGlu 5
|||||
DB 86 GGCTACTCGGTGAG 100

RESULT 17
US-09-063-733A-53/c
; Sequence 53, Application US/09063733A
; Patent No. 6372211
; GENERAL INFORMATION:
; APPLICANT: Isaac, Barbara G.
; APPLICANT: Greenplate, John T.

; APPLICANT: Purcell, John P.
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: INSECTS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold White & Durkee
; STREET: PO Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/063,733A
; FILING DATE: 21-APR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Patterson, Melinda L.
; REGISTRATION NUMBER: 33,062
; REFERENCE/DOCKET NUMBER: MOBT-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-787-1440
; INFORMATION FOR SEQ ID NO: 53:
; LENGTH: 193 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-063-733A-53

Alignment Scores:
Pred. No.: 394 Length: 193
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-063-733A-53 (1-193)

QY 1 GlyTyr***ValGlu 5
|||||
DB 147 GGCTACACGGTTGAG 133

RESULT 18
US-09-016-434-285/c
; Sequence 285, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/09/016.434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 285:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 202 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADNOT06
; CLONE: 1723064
US-09-016-434-285

```

```

Alignment Scores:
Pred. No.: 414 Length: 202
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

```

US-10-030-194A-6 (1-6) x US-09-016-434-285 (1-202)

```

Qy 1 GlyTyr***ValGlu 5
Db 86 GGCATACAGTGGAG 72

```

```

RESULT 19
US-09-513-999C-35499
; Sequence 35499, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 35499
; LENGTH: 206
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-35499

```

```

Alignment Scores:
Pred. No.: 423 Length: 206
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

```

US-10-030-194A-6 (1-6) x US-09-513-999C-35499 (1-206)

```

Qy 1 GlyTyr***ValGlu 5

```

```

Db 154 GCGTATGCTGTGAG 168

```

```

RESULT 20
US-09-134-000C-2517/c
; Sequence 2517, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2517
; LENGTH: 225
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2517

```

```

Alignment Scores:
Pred. No.: 466 Length: 225
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

```

US-10-030-194A-6 (1-6) x US-09-134-000C-2517 (1-225)

```

Qy 1 GlyTyr***ValGlu 5
Db 118 GGTATAGTGTGAA 104

```

```

RESULT 21
US-09-134-000C-1406
; Sequence 1406, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1406
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1406

```

```

Alignment Scores:
Pred. No.: 473 Length: 228
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

```

US-10-030-194A-6 (1-6) x US-09-134-000C-1406 (1-228)

```

Qy 1 GlyTyr***ValGlu 5
Db 131 GGTATTTCAGTGGAA 145

```

RESULT 22  
US-09-696-569-1  
; Sequence 1, Application US/09696569  
; Patent No. 6465215  
; GENERAL INFORMATION:  
; APPLICANT: Price, Jack  
; APPLICANT: Uwanogho, Dafe  
; TITLE OF INVENTION: Identification of Cells for Transplantation  
; FILE REFERENCE: GJE-36XCI  
; CURRENT APPLICATION NUMBER: US/09/696,569  
; CURRENT FILING DATE: 2000-10-25  
; PRIOR APPLICATION NUMBER: 9925210.8  
; PRIOR FILING DATE: 1999-10-25  
; PRIOR APPLICATION NUMBER: 60/170,692  
; PRIOR FILING DATE: 1999-12-14  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 273  
; TYPE: DNA  
; ORGANISM: Mouse  
US-09-696-569-1

Alignment Scores:  
Pred. No.: 575 Length: 273  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-696-569-1 (1-273)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 126 GGATATGCCGTGGAA 140

RESULT 23  
US-09-513-999C-28003  
; Sequence 28003, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 28003  
; LENGTH: 277  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 135  
; OTHER INFORMATION: n=a, g, c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 136  
; OTHER INFORMATION: k=g or t  
US-09-513-999C-28003

Alignment Scores:  
Pred. No.: 584 Length: 277  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-513-999C-28003 (1-277)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 105 GGCTACACATAGAA 119

RESULT 24  
US-09-513-999C-26667/c  
; Sequence 26667, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 26667  
; LENGTH: 285  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-513-999C-26667

Alignment Scores:  
Pred. No.: 603 Length: 285  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-513-999C-26667 (1-285)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 134 GGATATACCGTGAA 120

RESULT 25  
US-08-447-173A-59  
; Sequence 59, Application US/08447173A  
; Patent No. 5843730  
; GENERAL INFORMATION:  
; APPLICANT: Wain-Hobson, Simon  
; APPLICANT: Pezo, Valerie  
; TITLE OF INVENTION: Method For Hypermutagenesis Of  
; TITLE OF INVENTION: Nucleotides  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I. Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,173A  
; FILING DATE: 19-NAY-1995  
; CLASSIFICATION: 435

Alignment Scores:		605	288
Pred. No.:		22.00	Matches: 4
Score:		80.00%	Conservative: 0
Percent Similarity:		80.00%	Mismatches: 1
Best Local Similarity:		91.67%	Indels: 0
Query Match:		2	Gaps: 0
DB:			
US-10-030-194A-6 (1-6) x US-08-447-173A-60 (1-286)			
QY	1 GlyTyr***Valglu S		
Db	183 GGUACGCCGUCAG 197		
RESULT 27			
US-09-720-655B-3/c			
; Sequence 3, Application US/09720655B			
; Patent No. 6723521			
; GENERAL INFORMATION:			
; APPLICANT: YOSHIMOTO, MAKOTO			
; APPLICANT: YAZAKI, MADOKA			
; APPLICANT: MATSUMOTO, KAYO			
; APPLICANT: TAKAYAMA, KIYOSHI			
; APPLICANT: TSURITANI, KATSUKI			
; TITLE OF INVENTION: SUGAR TRANSPORTER			
; FILE REFERENCE: ASA-C034			
; CURRENT APPLICATION NUMBER: US/09/720,655B			
; CURRENT FILING DATE: 2001-02-28			
; PRIOR APPLICATION NUMBER: JP 10/187235			
; PRIOR FILING DATE: 1998-07-02			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: Patent In Ver. 2.1			
; SEQ ID NO 3			
; LENGTH: 292			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: modified_base			
; LOCATION: (36)			
; OTHER INFORMATION: a, c, g, t, unknown or other			
; FEATURE:			
; NAME/KEY: modified_base			
; LOCATION: (47)			
; OTHER INFORMATION: a, c, g, t, unknown or other			
; FEATURE:			
; NAME/KEY: modified_base			
; LOCATION: (251)			
; OTHER INFORMATION: a, c, g, t, unknown or other			
; FEATURE:			
; NAME/KEY: modified_base			
; LOCATION: (273)			
; OTHER INFORMATION: a, c, g, t, unknown or other			
US-09-720-655B-3			
Alignment Scores:			
Pred. No.:	619	Length:	292
Score:	22.00	Matches:	4
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	1
Query Match:	91.67%	Indels:	0
DB:	4	Gaps:	0
US-10-030-194A-6 (1-6) x US-09-720-655B-3 (1-292)			
QY	1 GlyTyr***Valglu S		
Db	173 GGTTCACAGTGGAG 159		
RESULT 28			
US-09-609-324A-7			
; Sequence 7, Application US/09609324A			
; Patent No. RE37582			
; GENERAL INFORMATION:			

; APPLICANT: CERRETTI, Douglas P.  
; TITLE OF INVENTION: CYTOKINE DESIGNATED LERK-6  
; FILE REFERENCE: A7772  
; CURRENT APPLICATION NUMBER: US/09/609,324A  
; CURRENT FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 08/920,440  
; PRIOR FILING DATE: 1997-08-29  
; PRIOR APPLICATION NUMBER: 08/538,709  
; PRIOR FILING DATE: 1995-10-03  
; PRIOR APPLICATION NUMBER: 08/318,393  
; PRIOR FILING DATE: 1994-10-05  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 314  
; TYPE: DNA  
; ORGANISM: LERK-6  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)..(313)  
US-09-609-324A-7

Alignment Scores: 314  
Pred. No.: 670 Length: 314  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 0  
Query Match: 91.67% Indels: 0  
DB: 1 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-609-324A-7 (1-314)

QY 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 32 GGCTACACGGTGGAG 46

## RESULT 29

US-08-920-440B-7  
; Sequence 7, Application US/08920440B  
; Patent No. 5919905  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated LERK-6  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: System 7.6  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,440B  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2826-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 314 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2..313  
US-08-920-440B-7

Alignment Scores: 314  
Pred. No.: 670 Length: 314  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x US-08-920-440B-7 (1-314)

QY 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 32 GGCTACACGGTGGAG 46

## RESULT 30

US-09-173-492-7  
; Sequence 7, Application US/09173492  
; Patent No. 6194172  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated LERK-6  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: System 7.6  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/173,492  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,440  
; FILING DATE: 29-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2826-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 314 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2..313  
US-09-173-492-7

Alignment Scores: 314  
Pred. No.: 670 Length: 314  
Score: 22.00 Matches: 4

Percent Similarity: 80.00%      Conservative: 0  
Best Local Similarity: 80.00%      Mismatches: 1  
Query Match: 91.67%              Indels: 0  
DB: 3                              Gaps: 0

US-10-030-194A-6 (1-6) x US-09-173-492-7 (1-314)

Qy 1 GlyTyr\*\*\*ValGlu 5  
||| ||| ||| ||| |||  
Db 32 GGCTACACGGTGGAG 46

RESULT 31

US-09-173-133-7  
; Sequence 7, Application US/09173133  
; Patent No. 6232447  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated LERK-6  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: System 7.6  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/173,133  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,440  
; FILING DATE: 29-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2826-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 314 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2..313  
US-09-173-133-7

Alignment Scores: 670      Length: 314  
Pred. No.: 22.00      Matches: 4  
Score: 80.00%      Conservative: 0  
Percent Similarity: 80.00%      Mismatches: 1  
Best Local Similarity: 80.00%      Indels: 0  
Query Match: 91.67%      Gaps: 0  
DB: 3

US-10-030-194A-6 (1-6) x US-09-173-133-7 (1-314)

Qy 1 GlyTyr\*\*\*ValGlu 5  
||| ||| ||| ||| |||  
Db 32 GGCTACACGGTGGAG 46

RESULT 32  
US-09-165-533-7  
; Sequence 7, Application US/09165533  
; Patent No. 6268482  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated as LERK-6  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/165,533  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/538,709  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Malaska, Stephen L.  
; REGISTRATION NUMBER: 32,655  
; REFERENCE/DOCKET NUMBER: 2826  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 314 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2..313  
US-09-165-533-7

Alignment Scores: 670      Length: 314  
Pred. No.: 22.00      Matches: 4  
Score: 80.00%      Conservative: 0  
Percent Similarity: 80.00%      Mismatches: 1  
Best Local Similarity: 80.00%      Indels: 0  
Query Match: 91.67%      Gaps: 0  
DB: 3

US-10-030-194A-6 (1-6) x US-09-165-533-7 (1-314)

Qy 1 GlyTyr\*\*\*ValGlu 5  
||| ||| ||| ||| |||  
Db 32 GGCTACACGGTGGAG 46

RESULT 33

US-09-580-236A-7  
; Sequence 7, Application US/09580236A  
; Patent No. 6472174  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated LERK-6  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street



```
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/580,236A
;   FILING DATE: 26-May-2000
;   CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Henry, Janis C.
;   REGISTRATION NUMBER: 34,347
;   REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 587-0430
;   TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 314 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 2..313
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-580-236A-7

Alignment Scores:
Pred. No.: 670 Length: 314
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-580-236A-7 (1-314)
QY 1 GlyTyr***ValGlu 5
DB 32 GGCTACACGGTGGAG 46

RESULT 34
US-09-323-873A-9
; Sequence 9, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 322

; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-323-873A-9

Alignment Scores:
Pred. No.: 689 Length: 322
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 3 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-323-873A-9 (1-322)
QY 1 GlyTyr***ValGlu 5
DB 224 GGTACAGTGTAGAA 238

RESULT 35
US-08-956-171E-4450
; Sequence 4450, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;   Gil H. Choi
;   Patrick S. Dillon
;   Craig A. Rosen
;   Steven C. Barash
;   Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 4450:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 326 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4450:
US-08-956-171E-4450

Alignment Scores:
Pred. No.: 698 Length: 326
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
```

Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0  
US-10-030-194A-6 (1-6) x US-08-956-171E-4450 (1-326)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 300 GGCTATACGGTAGAA 314  
RESULT 36  
US-08-781-986A-4450  
; Sequence 4450, Application US/08781986A  
; Patent No. 6737248  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781.986A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benson, Bob  
; REGISTRATION NUMBER: 30,446  
; REFERENCE/DOCKET NUMBER: PB248PP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 4450:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 326 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-781-986A-4450  
Alignment Scores:  
Pred. No.: 698 Length: 326  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0  
US-10-030-194A-6 (1-6) x US-08-781-986A-4450 (1-326)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 300 GGCTATACGGTAGAA 314  
RESULT 37  
US-08-929-856-65/c  
; Sequence 65, Application US/08929856  
; Patent No. 6136568  
; GENERAL INFORMATION:  
; APPLICANT: Hiatt, Andrew  
; APPLICANT: Rose, Floyd  
; TITLE OF INVENTION: DE NOVO POLYNUCLEOTIDE SYNTHESIS USING

; TITLE OF INVENTION: ROLLING TEMPLATES  
; NUMBER OF SEQUENCES: 190  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ &  
; ADDRESSEE: MENTILIK  
; STREET: 600 South, Avenue West  
; CITY: Westfield  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07090  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/929,856  
; FILING DATE: 15-SEP-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Foley, Shawn P.  
; REGISTRATION NUMBER: 33,071  
; REFERENCE/DOCKET NUMBER: ROSE 3.0-057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 908-654-5000  
; TELEFAX: 908-654-7866  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 342 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..342  
; US-08-929-856-65  
Alignment Scores:  
Pred. No.: 735 Length: 342  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 3 Gaps: 0  
US-10-030-194A-6 (1-6) x US-08-929-856-65 (1-342)  
Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 102 GGTTATACGTGTGAG 88  
RESULT 38  
US-08-929-856-188/c  
; Sequence 188, Application US/08929856  
; Patent No. 6136568  
; GENERAL INFORMATION:  
; APPLICANT: Hiatt, Andrew  
; APPLICANT: Rose, Floyd  
; TITLE OF INVENTION: DE NOVO POLYNUCLEOTIDE SYNTHESIS USING  
; TITLE OF INVENTION: ROLLING TEMPLATES  
; NUMBER OF SEQUENCES: 190  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ &  
; ADDRESSEE: MENTILIK  
; STREET: 600 South, Avenue West  
; CITY: Westfield  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07090  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

Qy 1 GlyTyr\*\*ValGlu 5  
|||  
Db 16 GGCTATTTCAGTGGAA 30

RESULT 42

US-09-030-607-215/c  
; Sequence 215, Application US/09030607  
; Patent No. 6262245  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO  
; NUMBER OF SEQUENCES: 224  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/030,607  
; FILING DATE: 25-FEB-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.427C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 215:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 366 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-030-607-215  
Alignment Scores: 792 Length: 366  
Pred. No.: 22.00 Matches: 4  
Score: 22.00  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 3 Gaps: 0  
US-10-030-194A-6 (1-6) x US-09-030-607-215 (1-366)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 326 GGTACAGTGTAGAA 312  
RESULT 43  
US-09-439-313-215/c  
; Sequence 215, Application US/09439313  
; Patent No. 6329505  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan Louise  
; APPLICANT: Jiang Yuqui  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Kalos, Michael  
; APPLICANT: Fanger, Gary  
; APPLICANT: Retter, Mark  
; APPLICANT: Solk, John  
; APPLICANT: Day, Craig  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; FILE REFERENCE: 210121.427C9

; CURRENT APPLICATION NUMBER: US/09/439,313  
; CURRENT FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 575  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 215  
; LENGTH: 366  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(366)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-439-313-215  
Alignment Scores: 792 Length: 366  
Pred. No.: 22.00 Matches: 4  
Score: 22.00  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 3 Gaps: 0  
US-10-030-194A-6 (1-6) x US-09-439-313-215 (1-366)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 326 GGTACAGTGTAGAA 312  
RESULT 44  
US-09-352-616A-215/c  
; Sequence 215, Application US/09352616A  
; Patent No. 6395278  
; GENERAL INFORMATION:  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Harlocker, Susan Louise  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Mitcham, Jennifer Lynn  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; FILE REFERENCE: 210121.427C8  
; CURRENT APPLICATION NUMBER: US/09/352,616A  
; CURRENT FILING DATE: 1999-07-13  
; NUMBER OF SEQ ID NOS: 472  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 215  
; LENGTH: 366  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(366)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-352-616A-215  
Alignment Scores: 792 Length: 366  
Pred. No.: 22.00 Matches: 4  
Score: 22.00  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 3 Gaps: 0  
US-10-030-194A-6 (1-6) x US-09-352-616A-215 (1-366)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 326 GGTACAGTGTAGAA 312  
RESULT 45  
US-09-232-149A-215/c  
; Sequence 215, Application US/09232149A  
; Patent No. 6465611

```

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C6
; CURRENT APPLICATION NUMBER: US/09/232,149A
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(366)
; OTHER INFORMATION: n = A,T,C or G
US-09-232-149A-215

Alignment Scores:
Pred. No.: 792 Length: 366
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-232-149A-215 (1-366)

Qy 1 GlyTyr***ValGlu 5
Db 326 GGTACAGTGTAGAA 312

RESULT 46
US-09-159-812-215/c
; Sequence 215, Application US/09159812A
; Patent No. 6613872
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun C.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.428C5
; CURRENT APPLICATION NUMBER: US/09/159,812A
; CURRENT FILING DATE: 1998-09-23
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(366)
; OTHER INFORMATION: n = A,T,C or G
US-09-159-812-215

Alignment Scores:
Pred. No.: 792 Length: 366
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-159-812-215 (1-366)

Qy 1 GlyTyr***ValGlu 5
Db 326 GGTACAGTGTAGAA 312

```

```

RESULT 47
US-09-636-215-215/c
; Sequence 215, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stoik, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(366)
; OTHER INFORMATION: n = A,T,C or G
US-09-636-215-215

Alignment Scores:
Pred. No.: 792 Length: 366
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-636-215-215 (1-366)

Qy 1 GlyTyr***ValGlu 5
Db 326 GGTACAGTGTAGAA 312

RESULT 48
US-09-685-166A-215/c
; Sequence 215, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stoik, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun

```

```

; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(366)
; OTHER INFORMATION: n = A,T,C or G
US-09-685-166A-215

Alignment Scores:
Pred. No.: 792 Length: 366
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-685-166A-215 (1-366)

QY 1 GlyTyr***ValGlu 5
DB 326 GGTACAGTGTAGAA 312

RESULT 49
US-09-115-453-215/c
; Sequence 215, Application US/09115453B
; Patent No. 6657056
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C4
; CURRENT APPLICATION NUMBER: US/09/115,453B
; CURRENT FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(366)
; OTHER INFORMATION: n = A,T,C or G
US-09-115-453-215

Alignment Scores:
Pred. No.: 792 Length: 366
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-115-453-215 (1-366)

QY 1 GlyTyr***ValGlu 5
DB 326 GGTACAGTGTAGAA 312

RESULT 50
US-09-688-489-215/c

```

```

; Sequence 215, Application US/09688489
; Patent No. 6664377
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun C.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427D2
; CURRENT APPLICATION NUMBER: US/09/688,489
; CURRENT FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(366)
; OTHER INFORMATION: n = A,T,C or G
US-09-688-489-215

Alignment Scores:
Pred. No.: 792 Length: 366
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-688-489-215 (1-366)

QY 1 GlyTyr***ValGlu 5
DB 326 GGTACAGTGTAGAA 312

RESULT 51
US-09-679-426-215/c
; Sequence 215, Application US/09679426
; Patent No. 6759515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun C.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C20
; CURRENT APPLICATION NUMBER: US/09/679,426
; CURRENT FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 895
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(366)

```

; OTHER INFORMATION: n = A,T,C or G  
US-09-679-426-215

Alignment Scores: 792 Length: 366  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 4

US-10-030-194A-6 (1-6) x US-09-679-426-215 (1-366)

Oy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 326 GGATACGTGTAGAA 312

## RESULT 52

US-08-454-557C-35  
; Sequence 35, Application US/08454557C  
; Patent No. 5830670  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; APPLICANT: Wands, Jack R.  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; TITLE OF INVENTION: of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,557C  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ludwig, Steven R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0609.3840003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 372 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
US-08-454-557C-35

Alignment Scores: 806 Length: 372  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 2

US-10-030-194A-6 (1-6) x US-08-454-557C-35 (1-372)

Oy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 187 GGATACGTGTAGAG 201

## RESULT 53

US-08-340-426D-35  
; Sequence 35, Application US/08340426D  
; Patent No. 5948634  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; APPLICANT: Wands, Jack R.  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; TITLE OF INVENTION: of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/340,426D  
; FILING DATE: 14-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ludwig, Steven R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0609.3840002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 372 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
US-08-340-426D-35

Alignment Scores: 806 Length: 372  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 2

US-10-030-194A-6 (1-6) x US-08-340-426D-35 (1-372)

Oy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 187 GGATACGTGTAGAG 201

## RESULT 54

US-08-450-673C-35  
; Sequence 35, Application US/08450673C  
; Patent No. 5948888  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; APPLICANT: Wands, Jack R.  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; TITLE OF INVENTION: of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
US-08-450-673C-35

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,673C  
FILING DATE: 30-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, Steven R.  
REGISTRATION NUMBER: 36,203  
REFERENCE/DOCKET NUMBER: 0609.3840004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
US-08-450-673C-35  
Alignment Scores:  
Pred. No.: 806 Length: 372  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
Gaps: 0  
US-10-030-194A-6 (1-6) x US-08-450-673C-35 (1-372)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 187 GGATATGCTGTAGAG 201  
RESULT 55  
PCT-US95-17111A-35  
Sequence 35, Application PC/TUS951711A  
GENERAL INFORMATION:  
APPLICANT: de la Monte, Suzanne  
APPLICANT: Wands, Jack R.  
TITLE OF INVENTION: Neural Thread Protein Gene Expression and  
TITLE OF INVENTION: Detection of Alzheimer's Disease  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/17111A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/340,426  
FILING DATE: 14-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, Steven R.  
REGISTRATION NUMBER: 36,203  
REFERENCE/DOCKET NUMBER: 0609.3840002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:

LENGTH: 372 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
PCT-US95-17111A-35  
Alignment Scores:  
Pred. No.: 806 Length: 372  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
Gaps: 0  
US-10-030-194A-6 (1-6) x PCT-US95-17111A-35 (1-372)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 187 GGATATGCTGTAGAG 201  
RESULT 56  
US-09-702-705-1768/c  
Sequence 1768, Application US/09702705  
Patent No. 6504010  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.478C14  
CURRENT APPLICATION NUMBER: US/09/702,705  
CURRENT FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 1833  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO 1768  
LENGTH: 384  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-702-705-1768  
Alignment Scores:  
Pred. No.: 834 Length: 384  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
Gaps: 0  
US-10-030-194A-6 (1-6) x US-09-702-705-1768 (1-384)  
QY 1 GlyTyr\*\*\*ValGlu 5  
Db 213 GGTATGCACTTGAG 199  
RESULT 57  
US-09-736-457-1768/c  
Sequence 1768, Application US/09736457  
Patent No. 6509448  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc



```
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1768
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1768

Alignment Scores:
Pred. No.: 834 Length: 384
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-736-457-1768 (1-384)

Qy 1 GlyTyr***ValGlu 5
Db 213 GGTATGCGAGTTGAG 199

RESULT 58
US-09-671-325-1768/c
; Sequence 1768, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1768
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-1768

Alignment Scores:
Pred. No.: 834 Length: 384
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-671-325-1768 (1-384)

Qy 1 GlyTyr***ValGlu 5
Db 213 GGTATGCGAGTTGAG 199

RESULT 59
US-09-658-824-1768/c
```

```
; Sequence 1768, Application US/09658824
; Patent No. 6748846
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C11
; CURRENT APPLICATION NUMBER: US/09/658,824
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 1788
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1768
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-658-824-1768

Alignment Scores:
Pred. No.: 834 Length: 384
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-658-824-1768 (1-384)

Qy 1 GlyTyr***ValGlu 5
Db 213 GGTATGCGAGTTGAG 199

RESULT 60
US-08-621-751A-11/c
; Sequence 11, Application US/08621751A
; Patent No. 5882644
; GENERAL INFORMATION:
; APPLICANT: Chang, Chung N.
; APPLICANT: Landolfi, Nicholas F.
; APPLICANT: Martin, Ulrich
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES SPECIFIC FOR THE
; TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR BETA RECEPTOR AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER LLP
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/621,751A
; FILING DATE: 22-MAR-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lehnhardt, Susan K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 321152000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 813-5600
```

Thu Nov 4 17:32:32 2004

```
;
; TELEFAX: (650) 494-0792
; TELEX: 706141 MRSN FOERS SFO
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 405 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..405
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 58..405
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: group(147..162, 205..255, 352..381)
; OTHER INFORMATION: /note= "Complementarity Determining
; OTHER INFORMATION: Regions(CDR-1, CDR-2 and CDR-3)"
US-08-621-751A-11

Alignment Scores:
Pred. No.: 884 Length: 405
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-194A-6 (1-6) x US-08-621-751A-11 (1-405)
QY 1 GlyTyr***ValGlu 5
Db 313 GGCTACTCAGTTGAA 299

RESULT 61
US-09-489-039A-1457
; Sequence 1457, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1457
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1457

Alignment Scores:
Pred. No.: 905 Length: 414
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-489-039A-1457 (1-414)
QY 1 GlyTyr***ValGlu 5
Db 330 GGCTATTCCGTGGAG 344

RESULT 62
US-09-976-594-512/c
; Sequence 512, Application US/09976594
; Patent No. 6673549
;
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 512
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 239996.2
US-09-976-594-512

Alignment Scores:
Pred. No.: 915 Length: 418
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-976-594-512 (1-418)
QY 1 GlyTyr***ValGlu 5
Db 36 GGGTACTCCGTGGAA 22

RESULT 63
US-09-134-000C-749
; Sequence 749, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 749
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-749

Alignment Scores:
Pred. No.: 920 Length: 420
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-134-000C-749 (1-420)
QY 1 GlyTyr***ValGlu 5
Db 331 GGTACACCGTGGAG 345

RESULT 64
US-09-621-976-17465
; Sequence 17465, Application US/09621976
; Patent No. 6639063
```

; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; PRIOR FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 17465  
; LENGTH: 424  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 422  
; OTHER INFORMATION: n=a, g, c or t  
US-09-621-976-17465

Alignment Scores:  
Pred. No.: 929 Length: 424  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-621-976-17465 (1-424)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||  
Db 136 GGATATGCTGTAGAG 150

RESULT 65  
US-09-621-976-10373/c  
; Sequence 10373, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 10373  
; LENGTH: 427  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-10373

Alignment Scores:  
Pred. No.: 936 Length: 427  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-621-976-10373 (1-427)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||  
Db 166 GGCTACTCTGTGAA 152

RESULT 66  
US-09-252-991A-16414  
; Sequence 16414, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 16414  
; LENGTH: 432  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16414

Alignment Scores:  
Pred. No.: 948 Length: 432  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-252-991A-16414 (1-432)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||  
Db 28 GGTATGCGATGGAA 42

RESULT 67  
US-08-956-171E-3176  
; Sequence 3176, Application US/08956171E  
; Patent No. 6593114  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956,171E  
; FILING DATE: 20-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,861  
; FILING DATE: January 5, 1996  
; APPLICATION NUMBER: 08/781,986  
; FILING DATE: January 3, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark J. Hyman  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB248P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (240) 314-1224  
; TELEFAX: (301) 309-8439  
; INFORMATION FOR SEQ ID NO: 3176:

SEQUENCE CHARACTERISTICS:  
LENGTH: 472 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 3176:  
US-08-956-171E-3176

Alignment Scores:  
Pred. No.: 1.04e+03 Length: 472  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-08-956-171E-3176 (1-472)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 283 GGTATTCTGTGAA 297

RESULT 68  
US-08-781-986A-3176  
; Sequence 3176, Application US/08781986A  
; Patent No. 6737248  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781,986A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benson, Bob  
; REGISTRATION NUMBER: 30,446  
; REFERENCE/DOCKET NUMBER: PB248PP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 3176:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 472 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-781-986A-3176

Alignment Scores:  
Pred. No.: 1.04e+03 Length: 472  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-08-781-986A-3176 (1-472)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 283 GGTATTCTGTGAA 297  
RESULT 69  
US-09-621-976-2058  
; Sequence 2058, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2058  
; LENGTH: 475  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 88..312  
; NAME/KEY: sig\_peptide  
; LOCATION: 88..198  
; OTHER INFORMATION: Von Heijne matrix  
; OTHER INFORMATION: score 3.5  
; OTHER INFORMATION: seq FFFISKLLPVFS/MM  
US-09-621-976-2058

Alignment Scores:  
Pred. No.: 1.05e+03 Length: 475  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-621-976-2058 (1-475)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 215 GGATATCTGTAGAG 229

RESULT 70  
US-09-621-976-14205/C  
; Sequence 14205, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 14205  
; LENGTH: 484  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-14205

Alignment Scores:  
Pred. No.: 1.07e+03 Length: 484  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0

DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-621-976-14205 (1-484)

Qy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 122 GGCTATACGGTAGAG 108

RESULT 71

US-09-252-991A-8516  
; Sequence 8516, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 8516  
; LENGTH: 486  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-8516

Alignment Scores:  
Pred. No.: 1.08e+03 Length: 486  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-252-991A-8516 (1-486)

Qy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 160 GGCTACTCGGTGGAA 174

RESULT 72

US-09-270-767-3772/c  
; Sequence 3772, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3772  
; LENGTH: 489  
; TYPE: DNA  
; ORGANISM: *Drosophila melanogaster*  
; FEATURE:  
; OTHER INFORMATION: n means any nucleotide

US-09-270-767-3772

Alignment Scores:  
Pred. No.: 1.09e+03 Length: 489  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-270-767-3772 (1-489)

Qy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 161 GGCTACGCTGCGAA 147

RESULT 73

US-09-270-767-19054/c  
; Sequence 19054, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 19054  
; LENGTH: 489  
; TYPE: DNA  
; ORGANISM: *Drosophila melanogaster*  
; FEATURE:  
; OTHER INFORMATION: n means any nucleotide

US-09-270-767-19054

Alignment Scores:  
Pred. No.: 1.09e+03 Length: 489  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-270-767-19054 (1-489)

Qy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 161 GGCTACGCTGCGAA 147

RESULT 74

US-09-134-000C-1338  
; Sequence 1338, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1338  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: *Enterococcus faecalis*

US-09-134-000C-1338

Alignment Scores:  
Pred. No.: 1.1e+03 Length: 495  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 4 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-134-000C-1338 (1-495)

Qy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 190 GGCTACGAGTTGAA 204

Thu Nov 4 17:32:32 2004

RESULT 75  
US-09-621-976-3143  
; Sequence 3143, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 3143  
; LENGTH: 513  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 185..346  
US-09-621-976-3143

Alignment Scores:                      Length:                      513  
Pred. No.:                      1.14e+03                      Matches:                      4  
Score:                      22.00                      Conservative:                      0  
Percent Similarity:                      80.00%                      Mismatches:                      1  
Best Local Similarity:                      80.00%                      Indels:                      0  
Query Match:                      91.67%                      Gaps:                      0  
DB:                      4

US-10-030-194A-6 (1-6) x US-09-621-976-3143 (1-513)

Qy                      1 GlyTyr\*\*\*ValGlu 5  
                         |||||                      |||||  
Db                      317 GGATACACAGTAGAA 331

Search completed: November 3, 2004, 17:56:49  
Job time : 97.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 3, 2004, 16:31:21 ; Search time 274 Seconds  
(without alignments)  
112.283 Million cell updates/sec

Title: US-10-030-194A-6

Perfect score: 24

Sequence: 1 GYXVEX 6

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/p/US10030194/runat\_01112004.184849\_29071/app.query.fasta\_1.398  
-DB=Published\_Applications\_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62  
-TRANS=human40.cdi -LIST=1000 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10030194 @CGN 1 1 1068 @runat\_01112004.184849\_29071  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubna/US10\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubna/US11\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubna/US60\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

\*

Result No.	Score	Query Match	Length	DB ID	Description	
1	22	91.7	20	17	US-10-619-739-321	
c	2	22	25	15	US-10-098-2638-30016	
	3	22	31	16	US-10-665-460A-153	
4	22	91.7	35	16	US-10-665-460A-154	
5	22	91.7	60	10	US-09-908-975-13740	
c	6	22	91.7	60	US-09-908-975-22204	
7	22	91.7	65	10	US-09-908-975-3559	
c	8	22	91.7	75	US-08-781-986A-3338	
c	9	22	91.7	75	US-10-329-624-3338	
c	10	22	91.7	83	US-09-983-965-5349	
c	11	22	91.7	88	US-10-029-386-21494	
	12	22	91.7	95	8	US-08-860-844-61
13	22	91.7	95	15	US-10-407-543-61	
14	22	91.7	100	9	US-09-999-672-14	
15	22	91.7	100	9	US-09-999-672-17	
16	22	91.7	100	13	US-10-040-863-14	
17	22	91.7	100	13	US-10-040-863-17	
c	18	22	91.7	123	US-10-424-599-119315	
c	19	22	91.7	124	US-10-029-386-17841	
20	22	91.7	126	17	US-10-794-929-96	
c	21	22	91.7	127	US-10-242-535A-58794	
c	22	22	91.7	127	US-10-424-599-38666	
c	23	22	91.7	127	US-10-085-783A-58794	
c	24	22	91.7	127	US-10-794-929-97	
c	25	22	91.7	127	US-10-794-929-98	
26	22	91.7	129	16	US-10-242-535A-17426	
27	22	91.7	129	16	US-10-424-599-19245	
c	28	22	91.7	129	US-10-085-783A-17426	
c	29	22	91.7	136	US-09-864-761-29192	
c	30	22	91.7	153	US-10-424-599-133516	
c	31	22	91.7	166	US-08-781-986A-4703	
32	22	91.7	166	16	US-10-329-624-4703	
c	33	22	91.7	182	US-10-424-599-93364	
c	34	22	91.7	189	US-09-864-761-21786	
c	35	22	91.7	193	US-10-005-530-53	
36	22	91.7	200	16	US-10-242-535A-58809	
37	22	91.7	200	16	US-10-085-783A-58809	
c	38	22	91.7	202	US-10-305-720-285	
c	39	22	91.7	205	US-10-424-599-26598	
c	40	22	91.7	207	US-10-437-963-4316	
c	41	22	91.7	219	US-10-437-963-71	
c	42	22	91.7	227	US-09-983-965-1009	
c	43	22	91.7	234	US-10-437-963-33060	
c	44	22	91.7	236	US-10-424-599-27759	
45	22	91.7	241	9	US-09-960-352-6378	
c	46	22	91.7	241	US-10-424-599-83707	
c	47	22	91.7	248	US-10-424-599-121558	
c	48	22	91.7	248	US-10-437-963-52740	
c	49	22	91.7	251	US-09-741-669-45	
c	50	22	91.7	251	US-10-282-122A-1445	
51	22	91.7	253	16	US-10-424-599-85528	
c	52	22	91.7	262	US-10-437-963-86065	
c	53	22	91.7	264	US-10-424-599-121254	
54	22	91.7	265	10	US-09-764-872-548	
55	22	91.7	267	10	US-09-764-872-37	
56	22	91.7	267	16	US-10-424-599-73754	
c	57	22	91.7	271	US-10-424-599-58619	
	58	22	91.7	273	14	US-10-140-463-1
59	22	91.7	278	16	US-10-424-599-78977	
c	60	22	91.7	286	US-09-741-669-46	
c	61	22	91.7	286	US-10-282-122A-1421	
c	62	22	91.7	289	US-10-242-535A-46051	
c	63	22	91.7	289	US-10-424-599-29094	
64	22	91.7	289	16	US-10-085-783A-46051	
65	22	91.7	298	16	US-10-424-599-18133	
66	22	91.7	300	14	US-10-076-555-87	
c	67	22	91.7	309	11	US-09-864-408A-1975
c	68	22	91.7	310	16	US-10-424-599-104737
69	22	91.7	316	17	US-10-437-963-28622	
70	22	91.7	318	11	US-09-864-408A-3483	















```
947 22 91.7 1759 14 US-10-176-993-151 Sequence 151, App
948 22 91.7 1759 14 US-10-184-658-151 Sequence 151, App
949 22 91.7 1759 14 US-10-176-991-151 Sequence 151, App
950 22 91.7 1759 14 US-10-173-695-151 Sequence 151, App
951 22 91.7 1759 14 US-10-173-697-151 Sequence 151, App
952 22 91.7 1759 14 US-10-173-705-151 Sequence 151, App
953 22 91.7 1759 14 US-10-174-576-151 Sequence 151, App
954 22 91.7 1759 14 US-10-174-585-151 Sequence 151, App
955 22 91.7 1759 14 US-10-174-586-151 Sequence 151, App
956 22 91.7 1759 14 US-10-175-747-151 Sequence 151, App
957 22 91.7 1759 14 US-10-176-481-151 Sequence 151, App
958 22 91.7 1759 14 US-10-176-485-151 Sequence 151, App
959 22 91.7 1759 14 US-10-176-487-151 Sequence 151, App
960 22 91.7 1759 14 US-10-176-493-151 Sequence 151, App
961 22 91.7 1759 14 US-10-176-756-151 Sequence 151, App
962 22 91.7 1759 14 US-10-176-911-151 Sequence 151, App
963 22 91.7 1759 14 US-10-176-919-151 Sequence 151, App
964 22 91.7 1759 14 US-10-176-925-151 Sequence 151, App
965 22 91.7 1759 14 US-10-176-978-151 Sequence 151, App
966 22 91.7 1759 14 US-10-179-510-151 Sequence 151, App
967 22 91.7 1759 14 US-10-180-543-151 Sequence 151, App
968 22 91.7 1759 14 US-10-180-544-151 Sequence 151, App
969 22 91.7 1759 14 US-10-180-546-151 Sequence 151, App
970 22 91.7 1759 14 US-10-180-547-151 Sequence 151, App
971 22 91.7 1759 14 US-10-180-549-151 Sequence 151, App
972 22 91.7 1759 14 US-10-180-555-151 Sequence 151, App
973 22 91.7 1759 14 US-10-180-559-151 Sequence 151, App
974 22 91.7 1759 14 US-10-181-000-151 Sequence 151, App
975 22 91.7 1759 14 US-10-183-010-151 Sequence 151, App
976 22 91.7 1759 14 US-10-183-012-151 Sequence 151, App
977 22 91.7 1759 14 US-10-184-614-151 Sequence 151, App
978 22 91.7 1759 14 US-10-184-623-151 Sequence 151, App
979 22 91.7 1759 14 US-10-184-635-151 Sequence 151, App
980 22 91.7 1759 14 US-10-184-637-151 Sequence 151, App
981 22 91.7 1759 14 US-10-184-646-151 Sequence 151, App
982 22 91.7 1759 14 US-10-184-647-151 Sequence 151, App
983 22 91.7 1759 14 US-10-184-652-151 Sequence 151, App
984 22 91.7 1759 14 US-10-187-594-151 Sequence 151, App
985 22 91.7 1759 14 US-10-187-596-151 Sequence 151, App
986 22 91.7 1759 14 US-10-187-745-151 Sequence 151, App
987 22 91.7 1759 14 US-10-187-885-151 Sequence 151, App
988 22 91.7 1759 14 US-10-187-886-151 Sequence 151, App
989 22 91.7 1759 14 US-10-199-464-151 Sequence 151, App
990 22 91.7 1759 14 US-10-196-756-151 Sequence 151, App
991 22 91.7 1759 14 US-10-176-751-151 Sequence 151, App
992 22 91.7 1759 14 US-10-176-760-151 Sequence 151, App
993 22 91.7 1759 14 US-10-176-990-151 Sequence 151, App
994 22 91.7 1759 14 US-10-180-541-151 Sequence 151, App
995 22 91.7 1759 14 US-10-180-542-151 Sequence 151, App
996 22 91.7 1759 14 US-10-180-548-151 Sequence 151, App
997 22 91.7 1759 14 US-10-180-551-151 Sequence 151, App
998 22 91.7 1759 14 US-10-180-998-151 Sequence 151, App
999 22 91.7 1759 14 US-10-180-999-151 Sequence 151, App
1000 22 91.7 1759 14 US-10-183-013-151 Sequence 151, App
```

```
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-619-739-321

Alignment Scores:
Pred. No.: 174 Length: 20
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-619-739-321 (1-20)

QY 1 GlyTyr***ValGlu 5
Db 6 GGATACACTGTTGAG 20

RESULT 2
US-10-098-263B-30016/c
; Sequence 30016, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 30016
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-30016

Alignment Scores:
Pred. No.: 221 Length: 25
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-098-263B-30016 (1-25)

QY 1 GlyTyr***ValGlu 5
Db 22 GGTACACTGTTGAG 8

RESULT 3
US-10-665-460A-153
; Sequence 153, Application US/10665460A
; Publication No. US20040096934A1
; GENERAL INFORMATION:
; APPLICANT: Prevysinet, Georges
; APPLICANT: Rang, Cecile
; APPLICANT: Frutos, Roger
; TITLE OF INVENTION: Pepsin-sensitive modified Bacillus thuringiensis insecticidal
; FILE REFERENCE: A35992-PCT-USA-A (072667.0191)
; CURRENT APPLICATION NUMBER: US/10/665,460A
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: PCT/FR02/00772
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: FR 01/03691
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 160
```

ALIGNMENTS

```
RESULT 1
US-10-619-739-321
; Sequence 321, Application US/10619739
; Publication No. US20040175719A1
; GENERAL INFORMATION:
; APPLICANT: Christians, Frederick C.
; TITLE OF INVENTION: Synthetic Tag Genes
; FILE REFERENCE: 3502.1
; CURRENT APPLICATION NUMBER: US/10/619,739
; CURRENT FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: 60/395,530
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 2068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 321
```

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 153  
; LENGTH: 31  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial sequence description:  
; OTHER INFORMATION: oligonucleotide 112  
US-10-665-460A-153

Alignment Scores:  
Pred. No.: 278 Length: 31  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-665-460A-153 (1-31)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 15 GGTACTCCGGTGAG 29

## RESULT 4

US-10-665-460A-154  
; Sequence 154, Application US/10665460A  
; Publication No. US20040096934A1  
; GENERAL INFORMATION:  
; APPLICANT: Freyessinet, Georges  
; APPLICANT: Rang, Cecile  
; APPLICANT: Frutos, Roger  
; TITLE OF INVENTION: Pepsin-sensitive modified Bacillus thuringiensis insecticidal  
; TITLE OF INVENTION: toxin  
; FILE REFERENCE: A35992-PCT-USA-A (072667.0191)  
; CURRENT APPLICATION NUMBER: US/10/665.460A  
; CURRENT FILING DATE: 2003-09-19  
; PRIOR APPLICATION NUMBER: PCT/FR02/00772  
; PRIOR FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: FR 01/03691  
; PRIOR FILING DATE: 2001-03-19  
; NUMBER OF SEQ ID NOS: 160  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 154  
; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial sequence description:  
; OTHER INFORMATION: oligonucleotide 113  
US-10-665-460A-154

Alignment Scores:  
Pred. No.: 317 Length: 35  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-665-460A-154 (1-35)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 5 GGTACTCCGGTGAG 19

## RESULT 5

US-09-908-975-13740  
; Sequence 13740, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi  
; APPLICANT: WASSERMAN, Alon

; APPLICANT: MINTZ, Eli  
; APPLICANT: FAIGLER, Simchon  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME  
; FILE REFERENCE: 36688-0005  
; CURRENT APPLICATION NUMBER: US/09/908,975  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/287,724  
; PRIOR FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: US 60/221,607  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 32337  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13740  
; LENGTH: 60  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-908-975-13740

Alignment Scores:  
Pred. No.: 564 Length: 60  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 10 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-908-975-13740 (1-60)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 46 GGTATGCACTTGAG 60

## RESULT 6

US-09-908-975-22204/C  
; Sequence 22204, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi  
; APPLICANT: WASSERMAN, Alon  
; APPLICANT: MINTZ, Eli  
; APPLICANT: FAIGLER, Simchon  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME  
; FILE REFERENCE: 36688-0005  
; CURRENT APPLICATION NUMBER: US/09/908,975  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/287,724  
; PRIOR FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: US 60/221,607  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 32337  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 22204  
; LENGTH: 60  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-908-975-22204

Alignment Scores:  
Pred. No.: 564 Length: 60  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 10 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-908-975-22204 (1-60)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||

```

; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446

```



```
;
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3338:
US-10-329-624-3338
Alignment Scores:
Pred. No.: 715 Length: 75
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-329-624-3338 (1-75)
Oy 1 GlyTyr***ValGlu 5
Db 23 GGCTATACGGTAGAA 9

RESULT 10
US-09-983-965-5349/c
; Sequence 5349, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 5349
; LENGTH: 83
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 41-LTB34-068-Q1-E1-C2
US-09-983-965-5349
Alignment Scores:
Pred. No.: 797 Length: 83
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-983-965-5349 (1-83)
Oy 1 GlyTyr***ValGlu 5
Db 72 GGTTATCCGTTGAG 58

RESULT 11
US-10-029-386-21494/c
; Sequence 21494, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21494
; LENGTH: 88
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR14.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: EST HUMAN HIT: BF698740.1. EVALUE 1.00e-01
; OTHER INFORMATION: NT HIT: gii4772816, EVALUE 1.10e+00
US-10-029-386-21494
Alignment Scores:
Pred. No.: 849 Length: 88
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 15 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-029-386-21494 (1-88)
Oy 1 GlyTyr***ValGlu 5
Db 31 GGATACACAGTGGAG 17

RESULT 12
US-08-860-844-61
; Sequence 61, Application US/08860844
; Publication No. US20030104361A1
; GENERAL INFORMATION:
; APPLICANT: Weininger, Susan
; APPLICANT: Weininger, Arthur M
; TITLE OF INVENTION: METHOD OF DETECTION OF NUCLEIC ACIDS WITH A
; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st St., Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/860,844
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/353,476
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: GP-100C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
```

MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-860-844-61

Alignment Scores:

Pred. No.: 921 Length: 95  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 8 Gaps: 0

US-10-030-194A-6 (1-6) x US-08-860-844-61 (1-95)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 20 GGATATACAGTGGAA 34

RESULT 13

US-10-407-543-61  
Sequence 61, Application US/10407543  
Publication No. US20030175789A1

GENERAL INFORMATION:

APPLICANT: Weininger, Susan  
TITLE OF INVENTION: METHOD OF DETECTION OF NUCLEIC ACIDS WITH A  
SPECIFIC SEQUENCE COMPOSITION

NUMBER OF SEQUENCES: 118

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik & Saliwanchik  
STREET: 2421 N.W. 41st St., Suite A-1  
CITY: Gainesville  
STATE: Florida  
COUNTRY: USA  
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/407,543  
FILING DATE: 03-APR-2003

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/860,844  
FILING DATE: 09-JUN-1997  
APPLICATION NUMBER: 08/353,476  
FILING DATE: 09-DEC-1994

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 35,746  
REFERENCE/DOCKET NUMBER: GP-100C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (352) 375-8100  
TELEFAX: (352) 372-5800

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:  
LENGTH: 95 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO

ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 61:

US-10-407-543-61  
Alignment Scores:  
Pred. No.: 921 Length: 95  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0

Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 15 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-407-543-61 (1-95)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 20 GGATATACAGTGGAA 34

RESULT 14

US-09-999-672-14  
Sequence 14, Application US/09999672  
Patent No. US20020127655A1

GENERAL INFORMATION:

APPLICANT: Eric H. Holmes et al.  
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
TITLE OF INVENTION: THEREOF

FILE REFERENCE: 8511-029

CURRENT APPLICATION NUMBER: US/09/999,672

CURRENT FILING DATE: 2001-10-31

PRIOR APPLICATION NUMBER: US/09/298,886

PRIOR FILING DATE: 1999-04-26

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 14

LENGTH: 100

TYPE: DNA

ORGANISM: Homo sapiens

US-09-999-672-14

Alignment Scores:

Pred. No.: 973 Length: 100  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-999-672-14 (1-100)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 11 GGCTACACCGTGGAA 25

RESULT 15

US-09-999-672-17

Sequence 17, Application US/09999672

Patent No. US20020127655A1

GENERAL INFORMATION:

APPLICANT: Eric H. Holmes et al.

TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
TITLE OF INVENTION: THEREOF

FILE REFERENCE: 8511-029

CURRENT APPLICATION NUMBER: US/09/999,672

CURRENT FILING DATE: 2001-10-31

PRIOR APPLICATION NUMBER: US/09/298,886

PRIOR FILING DATE: 1999-04-26

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 17

LENGTH: 100

TYPE: DNA

ORGANISM: Homo sapiens

US-09-999-672-17

Alignment Scores:

Pred. No.: 973 Length: 100  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0  
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-999-672-17 (1-100)

QY 1 GLYTYR\*\*\*ValGlu 5  
DB 11 GGCTACACCGTGGAA 25

RESULT 16  
US-10-040-863-14  
; Sequence 14, Application US/10040863  
; Publication No. US20020137165A1  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; FILE REFERENCE: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; TITLE OF INVENTION: THEREOF  
; CURRENT APPLICATION NUMBER: US/10/040,863  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 09/298,886  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 100  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-040-863-14

Alignment Scores:  
Pred. No.: 973 Length: 100  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 13 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-040-863-14 (1-100)

QY 1 GLYTYR\*\*\*ValGlu 5  
DB 11 GGCTACACCGTGGAA 25

RESULT 17  
US-10-040-863-17  
; Sequence 17, Application US/10040863  
; Publication No. US20020137165A1  
; GENERAL INFORMATION:  
; APPLICANT: Eric H. Holmes et al.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE  
; FILE REFERENCE: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES  
; TITLE OF INVENTION: THEREOF  
; CURRENT APPLICATION NUMBER: US/10/040,863  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 09/298,886  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 100  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-040-863-17

Alignment Scores:  
Pred. No.: 973 Length: 100  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 0  
Query Match: 91.67% Indels: 0  
DB: 13 Gaps: 0

DB: 13 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-040-863-17 (1-100)

QY 1 GLYTYR\*\*\*ValGlu 5  
DB 11 GGCTACACCGTGGAA 25

RESULT 18  
US-10-424-599-119315/c  
; Sequence 119315, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement  
; TITLE OF INVENTION: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 119315  
; LENGTH: 123  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_7874C.1  
US-10-424-599-119315

Alignment Scores:  
Pred. No.: 1,21e+03 Length: 123  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-119315 (1-123)

QY 1 GLYTYR\*\*\*ValGlu 5  
DB 97 GGTTATGCTGTGAA 83

RESULT 19  
US-10-029-386-17841/c  
; Sequence 17841, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AECOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 17841  
; LENGTH: 124  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL136122.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.63  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.46  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.5  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.81  
; OTHER INFORMATION: SWISSPROT HIT: P39120, EVALUE 7.90e+00  
US-10-029-386-17841

Alignment Scores: 124  
 Pred. No.: 1.22e+03  
 Score: 22.00  
 Percent Similarity: 80.00%  
 Best Local Similarity: 80.00%  
 Query Match: 91.67%  
 DB: 15  
 Gaps: 0  
 Indels: 0  
 Mismatches: 1  
 Conservatives: 0  
 Matches: 4

US-10-030-194A-6 (1-6) x US-10-029-386-17841 (1-124)

QY 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 DB 43 GGCTACACGGTTGAA 29

RESULT 20

US-10-794-929-96  
 ; Sequence 96, Application US/10794929  
 ; Publication No. US20040166526A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMNISCIENCE PHARMACEUTICALS  
 ; APPLICANT: Chaparian, Michael  
 ; APPLICANT: Zheng, Shu-Xian  
 ; TITLE OF INVENTION: GENE CLONING  
 ; FILE REFERENCE: 1002.00011  
 ; CURRENT APPLICATION NUMBER: US/10/794,929  
 ; CURRENT FILING DATE: 2004-03-05  
 ; PRIOR APPLICATION NUMBER: 10/049,994  
 ; PRIOR FILING DATE: 2002-02-18  
 ; PRIOR APPLICATION NUMBER: PCT/US00/22743  
 ; PRIOR FILING DATE: 2000-08-18  
 ; PRIOR APPLICATION NUMBER: 60/149,788  
 ; PRIOR FILING DATE: 1999-08-19  
 ; PRIOR APPLICATION NUMBER: 60/149,822  
 ; PRIOR FILING DATE: 1999-08-19  
 ; NUMBER OF SEQ ID NOS: 141  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 96  
 ; LENGTH: 126  
 ; TYPE: DNA  
 ; ORGANISM: artificial sequence  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)..(126)  
 ; OTHER INFORMATION: probe  
 US-10-794-929-96

Alignment Scores: 126  
 Pred. No.: 1.25e+03  
 Score: 22.00  
 Percent Similarity: 80.00%  
 Best Local Similarity: 80.00%  
 Query Match: 91.67%  
 DB: 17  
 Gaps: 0  
 Indels: 0  
 Mismatches: 1  
 Conservatives: 0  
 Matches: 4

US-10-030-194A-6 (1-6) x US-10-794-929-96 (1-126)

QY 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 DB 78 GGCTATGCCGTCGAG 92

RESULT 21

US-10-242-535A-58794/c  
 ; Sequence 58794, Application US/10242535A  
 ; Publication No. US20040013663A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ChondroGene Inc.  
 ; APPLICANT: Liew, C.C.  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
 ; FILE REFERENCE: 4231/2005  
 ; CURRENT APPLICATION NUMBER: US/10/242,535A  
 ; CURRENT FILING DATE: 2002-09-12  
 ; PRIOR APPLICATION NUMBER: US 10/085,783

; PRIOR FILING DATE: 2002-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/305,340  
 ; PRIOR FILING DATE: 2001-07-13  
 ; PRIOR APPLICATION NUMBER: US 60/275,017  
 ; PRIOR FILING DATE: 2001-03-12  
 ; PRIOR APPLICATION NUMBER: US 60/271,955  
 ; PRIOR FILING DATE: 2001-02-28  
 ; NUMBER OF SEQ ID NOS: 58994  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 58794  
 ; LENGTH: 127  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-10-242-535A-58794

Alignment Scores: 127  
 Pred. No.: 1.26e+03  
 Score: 22.00  
 Percent Similarity: 80.00%  
 Best Local Similarity: 80.00%  
 Query Match: 91.67%  
 DB: 16  
 Gaps: 0  
 Indels: 0  
 Mismatches: 1  
 Conservatives: 0  
 Matches: 4

US-10-030-194A-6 (1-6) x US-10-242-535A-58794 (1-127)

QY 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 DB 25 GGCTATAGTGTAGAG 11

RESULT 22

US-10-424-599-38666/c  
 ; Sequence 38666, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 38666  
 ; LENGTH: 127  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_134916C.1  
 US-10-424-599-38666

Alignment Scores: 127  
 Pred. No.: 1.26e+03  
 Score: 22.00  
 Percent Similarity: 80.00%  
 Best Local Similarity: 80.00%  
 Query Match: 91.67%  
 DB: 16  
 Gaps: 0  
 Indels: 0  
 Mismatches: 1  
 Conservatives: 0  
 Matches: 4

US-10-030-194A-6 (1-6) x US-10-424-599-38666 (1-127)

QY 1 GlyTyr\*\*\*ValGlu 5  
 |||||  
 DB 48 GGATATACGGTGGAA 34

RESULT 23

US-10-085-783A-58794/c  
 ; Sequence 58794, Application US/10085783A  
 ; Publication No. US20040037841A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ChondroGene Inc.  
 ; APPLICANT: Liew, C.C.

; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis

; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 58794  
; LENGTH: 127  
; TYPE: DNA

; ORGANISM: Human  
US-10-085-783A-58794

Alignment Scores:  
Pred. No.: 1.26e+03 Length: 127  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-085-783A-58794 (1-127)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 25 GGCTATAGTGTAGAG 11

## RESULT 24

; Sequence 97, Application US/10794929  
; Publication No. US20040166526A1  
; GENERAL INFORMATION:  
; APPLICANT: OMNISCIENCE PHARMACEUTICALS  
; APPLICANT: Chaparian, Michael  
; APPLICANT: Zheng, Shu-Xian  
; TITLE OF INVENTION: GENE CLONING  
; FILE REFERENCE: 1002.00011  
; CURRENT APPLICATION NUMBER: US/10/794,929  
; CURRENT FILING DATE: 2004-03-05  
; PRIOR APPLICATION NUMBER: 10/049,994  
; PRIOR FILING DATE: 2002-02-18  
; PRIOR APPLICATION NUMBER: PCT/US00/22743  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: 60/149,788  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: 60/149,822  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 141  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 97  
; LENGTH: 127  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(127)  
; OTHER INFORMATION: probe

Alignment Scores:  
Pred. No.: 1.26e+03 Length: 127  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 17 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-794-929-97 (1-127)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 49 GGCTATGCCGTCGAG 35

## RESULT 25

US-10-794-929-98/c  
; Sequence 98, Application US/10794929  
; Publication No. US20040166526A1  
; GENERAL INFORMATION:  
; APPLICANT: OMNISCIENCE PHARMACEUTICALS  
; APPLICANT: Chaparian, Michael  
; APPLICANT: Zheng, Shu-Xian  
; TITLE OF INVENTION: GENE CLONING  
; FILE REFERENCE: 1002.00011  
; CURRENT APPLICATION NUMBER: US/10/794,929  
; CURRENT FILING DATE: 2004-03-05  
; PRIOR APPLICATION NUMBER: 10/049,994  
; PRIOR FILING DATE: 2002-02-18  
; PRIOR APPLICATION NUMBER: PCT/US00/22743  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: 60/149,788  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: 60/149,822  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 141  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 98  
; LENGTH: 127  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(127)  
; OTHER INFORMATION: probe

US-10-794-929-98

Alignment Scores:  
Pred. No.: 1.26e+03 Length: 127  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 17 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-794-929-98 (1-127)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 49 GGCTATGCCGTCGAG 35

## RESULT 26

US-10-242-535A-17426  
; Sequence 17426, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liaw, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 17426

LENGTH: 129  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (33)..(33)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (37)..(37)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (65)..(65)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (68)..(68)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-242-535A-17426

Alignment Scores:  
Pred. No.: 1.28e+03 Length: 129  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-242-535A-17426 (1-129)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 103 GGCTATGCGGTCGAG 117

RESULT 27  
US-10-424-599-19245  
; Sequence 19245, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(S3223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 19245  
; LENGTH: 129  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_117382C.1  
US-10-424-599-19245

Alignment Scores:  
Pred. No.: 1.28e+03 Length: 129  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-19245 (1-129)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 113 GGGTACGTGTGGAG 127

RESULT 28

US-10-085-783A-17426  
; Sequence 17426, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liaw, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 17426  
; LENGTH: 129  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (33)..(33)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (37)..(37)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (65)..(65)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (68)..(68)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-085-783A-17426

Alignment Scores:  
Pred. No.: 1.28e+03 Length: 129  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-085-783A-17426 (1-129)

QY 1 GlyTyr\*\*\*ValGlu 5  
DB 103 GGCTATGCGGTCGAG 117

RESULT 29  
US-09-864-761-29192/c  
; Sequence 29192, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366

```
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29192
; LENGTH: 136
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP00431.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.57
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.52
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.63
; OTHER INFORMATION: NT HIT: M18095.1, EVALUE 4.70e+00
; OTHER INFORMATION: SWISSPROT HIT: P11253, EVALUE 4.10e+00
US-09-864-761-29192

Alignment Scores:
Pred. No.: 1.35e+03 Length: 136
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-864-761-29192 (1-136)

Oy 1 GlyTyr***ValGlu 5
||| ||| ||| |||
Db 127 GGTATTCCGTGGAA 113

RESULT 30
US-10-424-599-133516/c
; Sequence 133516, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 133516
; LENGTH: 153
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91570C.1
US-10-424-599-133516

Alignment Scores:
Pred. No.: 1.53e+03 Length: 153
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-133516 (1-153)

Oy 1 GlyTyr***ValGlu 5
||| ||| ||| |||
Db 31 GGTACACATGGAG 17

RESULT 31
US-08-781-986A-4703
; Sequence 4703, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 4703:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-4703

Alignment Scores:
Pred. No.: 1.67e+03 Length: 166
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
```

Query Match: 91.67% Indels: 0  
DB: 8 Gaps: 0  
US-10-030-194A-6 (1-6) x US-08-781-986A-4703 (1-166)  
QY 1 GLYTYR\*\*\*ValGlu 5  
DB 13 GGTATTCTGTGTGAA 27  
RESULT 32  
US-10-329-624-4703  
; Sequence 4703, Application US/10329624  
; Publication No. US20040043037A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/329,624  
; FILING DATE: 27-Dec-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/956,171  
; FILING DATE: October 20, 1997  
; APPLICATION NUMBER: 60/009,861  
; FILING DATE: January 5, 1996  
; APPLICATION NUMBER: 08/781,986  
; FILING DATE: January 3, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark J. Hyman  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB248P1D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (240) 314-1224  
; TELEFAX: (301) 309-8439  
; INFORMATION FOR SEQ ID NO: 4703:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 166 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4703:  
US-10-329-624-4703  
Alignment Scores: 1.67e+03 Length: 166  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 16  
US-10-030-194A-6 (1-6) x US-10-329-624-4703 (1-166)  
QY 1 GLYTYR\*\*\*ValGlu 5  
DB 13 GGTATTCTGTGTGAA 27

RESULT 33  
US-10-424-599-93364/c  
; Sequence 93364, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 93364  
; LENGTH: 182  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55321C.1  
US-10-424-599-93364  
Alignment Scores: 1.85e+03 Length: 182  
Pred. No.: 22.00 Matches: 4  
Score: 80.00% Conservative: 0  
Percent Similarity: 80.00% Mismatches: 1  
Best Local Similarity: 80.00% Indels: 0  
Query Match: 91.67% Gaps: 0  
DB: 16  
US-10-030-194A-6 (1-6) x US-10-424-599-93364 (1-182)  
QY 1 GLYTYR\*\*\*ValGlu 5  
DB 147 GGGTATGCCGTGGAA 133  
RESULT 34  
US-09-864-761-21786/c  
; Sequence 21786, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aesomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668



;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 21786  
;; LENGTH: 189  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AC009576.2  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.93  
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.87  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 10  
;; OTHER INFORMATION: EST HUMAN HIT: BE617547.1, EVALUE 2.60e+00  
;; OTHER INFORMATION: SWISSPROT HIT: Q58850, EVALUE 2.90e+00  
;; OTHER INFORMATION: NT HIT: U63374.1, EVALUE 9.50e-02  
;; NAME/KEY: unsure  
;; LOCATION: 182  
US-09-864-761-21786

Alignment Scores:  
Pred. No.: 1.92e+03 Length: 189  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-864-761-21786 (1-189)

Qy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 138 GGCTATGCTGTAGAA 124

## RESULT 35

US-10-005-530-53/c  
; Sequence 53, Application US/10005530  
; Publication No. US20030026795A1  
; GENERAL INFORMATION:  
; APPLICANT: Isaac, Barbara G.  
; APPLICANT: Greenplate, John T.  
; APPLICANT: Purcell, John P.  
; APPLICANT: Romano, Charles P.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING INSECTS  
; FILE REFERENCE: 11899.0022.DVUS01 (MOBT:022--2)  
; CURRENT APPLICATION NUMBER: US/10/005,530  
; PRIOR FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: 09/063,733  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/044,504  
; PRIOR FILING DATE: 1997-04-21  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 53  
; LENGTH: 193  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Polynucleotide

US-10-005-530-53

Alignment Scores:  
Pred. No.: 1.97e+03 Length: 193  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 14 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-005-530-53 (1-193)

Qy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 147 GGCTACACGGTTGAG 133

## RESULT 36

US-10-242-535A-58809  
; Sequence 58809, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 58809  
; LENGTH: 200  
; TYPE: DNA  
; ORGANISM: Human

US-10-242-535A-58809

Alignment Scores:  
Pred. No.: 2.04e+03 Length: 200  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-242-535A-58809 (1-200)

Qy 1 GlyTyr\*\*\*ValGlu 5  
| | | | |  
Db 63 GGATATTCTGTAGAA 77

## RESULT 37

US-10-085-783A-58809  
; Sequence 58809, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955

;  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 58809  
; LENGTH: 200  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-58809

Alignment Scores:  
Pred. No.: 2.04e+03 Length: 200  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-085-783A-58809 (1-200)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 63 GGATATTCTGTAGAA 77

RESULT 38  
US-10-305-720-285/c  
; Sequence 285, Application US/10305720  
; Publication No. US20040010136A1  
; GENERAL INFORMATION:  
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.  
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression  
; FILE REFERENCE: PA-0002-1 CON  
; CURRENT APPLICATION NUMBER: US/10/305,720  
; CURRENT FILING DATE: 2002-11-26  
; PRIOR APPLICATION NUMBER: 09/016,434  
; PRIOR FILING DATE: 1998-01-30  
; NUMBER OF SEQ ID NOS: 1490  
; SOFTWARE: PERL Program  
; SEQ ID NO 285  
; LENGTH: 202  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040010136A1 1723064  
US-10-305-720-285

Alignment Scores:  
Pred. No.: 2.06e+03 Length: 202  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-305-720-285 (1-202)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 86 GGCTATACAGTGGAG 72

RESULT 39  
US-10-424-599-26598/c  
; Sequence 26598, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599

;  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 26598  
; LENGTH: 205  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_124018C.1  
US-10-424-599-26598

Alignment Scores:  
Pred. No.: 2.1e+03 Length: 205  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-26598 (1-205)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 98 GGTTATGCTGTTGAA 84

RESULT 40  
US-10-437-963-4316  
; Sequence 4316, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 4316  
; LENGTH: 207  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_11210C.1  
US-10-437-963-4316

Alignment Scores:  
Pred. No.: 2.12e+03 Length: 207  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 17 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-437-963-4316 (1-207)

Qy 1 GlyTyr\*\*\*ValGlu 5  
Db 185 GGTTACGCGTGGAG 199

RESULT 41  
US-10-437-963-71/c  
; Sequence 71, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 71
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_100066C.1
US-10-437-963-71

Alignment Scores:
Pred. No.: 2.25e+03      Length: 219
Score: 22.00           Matches: 4
Percent Similarity: 80.00%      Conservative: 0
Best Local Similarity: 80.00%    Mismatches: 1
Query Match: 91.67%           Indels: 0
DB: 17                   Gaps: 0

US-10-030-194A-6 (1-6) x US-10-437-963-71 (1-219)

Qy 1 GlyTyr***ValGlu 5
Db 125 GGGTATGCAGTAGAG 111

RESULT 42
US-09-983-965-1009/c
; Sequence 1009, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 1009
; LENGTH: 227
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 50-LIB188-008-Q1-E1-E6
US-09-983-965-1009

Alignment Scores:
Pred. No.: 2.34e+03      Length: 227
Score: 22.00           Matches: 4
Percent Similarity: 80.00%      Conservative: 0
Best Local Similarity: 80.00%    Mismatches: 1
Query Match: 91.67%           Indels: 0
DB: 9                   Gaps: 0

US-10-030-194A-6 (1-6) x US-09-983-965-1009 (1-227)

Qy 1 GlyTyr***ValGlu 5
Db 213 GGATATACAGTGGAG 199
```

```
RESULT 43
US-10-437-963-33060/c
; Sequence 33060, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 33060
; LENGTH: 234
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37209C.1
US-10-437-963-33060

Alignment Scores:
Pred. No.: 2.42e+03      Length: 234
Score: 22.00           Matches: 4
Percent Similarity: 80.00%      Conservative: 0
Best Local Similarity: 80.00%    Mismatches: 1
Query Match: 91.67%           Indels: 0
DB: 17                   Gaps: 0

US-10-030-194A-6 (1-6) x US-10-437-963-33060 (1-234)

Qy 1 GlyTyr***ValGlu 5
Db 176 GGGTACGCAGTAGAG 162

RESULT 44
US-10-424-599-27759
; Sequence 27759, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 27759
; LENGTH: 236
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125068C.1
US-10-424-599-27759

Alignment Scores:
Pred. No.: 2.44e+03      Length: 236
Score: 22.00           Matches: 4
Percent Similarity: 80.00%      Conservative: 0
Best Local Similarity: 80.00%    Mismatches: 1
Query Match: 91.67%           Indels: 0
DB: 16                   Gaps: 0
```

```
US-10-030-194A-6 (1-6) x US-10-424-599-27759 (1-236)
QY 1 GlyTyr***ValGlu 5
Db 16 GGATATTCGGTTGAA 30

RESULT 45
US-09-960-352-6378
; Sequence 6378, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6378
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 27-LIB34-060-Q1-E1-G3
US-09-960-352-6378

Alignment Scores:
Pred. No.: 2.49e+03 Length: 241
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-960-352-6378 (1-241)
QY 1 GlyTyr***ValGlu 5
Db 123 GGATATGCAGTAGAG 137

RESULT 46
US-10-424-599-83707/c
; Sequence 83707, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 83707
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(241)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_46602C.1
US-10-424-599-83707

Alignment Scores:
Pred. No.: 2.49e+03 Length: 241
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0

US-10-030-194A-6 (1-6) x US-10-424-599-83707 (1-241)
QY 1 GlyTyr***ValGlu 5
Db 164 GGTATGCTGTGAA 150

RESULT 47
US-10-424-599-121558/c
; Sequence 121558, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 121558
; LENGTH: 248
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80774C.1
US-10-424-599-121558

Alignment Scores:
Pred. No.: 2.57e+03 Length: 248
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-121558 (1-248)
QY 1 GlyTyr***ValGlu 5
Db 164 GGTATGCTGTGAA 150

RESULT 48
US-10-437-963-52740/c
; Sequence 52740, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 52740
; LENGTH: 248
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55007C.1
US-10-437-963-52740
```

Alignment Scores: Length: 248  
Pred. No.: 2.57e+03  
Score: 22.00  
Percent Similarity: 80.00%  
Best Local Similarity: 80.00%  
Query Match: 91.67%  
DB: 17

US-10-030-194A-6 (1-6) x US-10-437-963-52740 (1-248)

Oy 1 GlyTyr\*\*\*ValGlu 5

Db 205 GGTATACCTGTGAA 191

RESULT 49

US-09-741-669-45/c  
; Sequence 45, Application US/09741669  
; Patent No. US2002022718A1  
; GENERAL INFORMATION:  
; APPLICANT: Forsyth, R. Allyn  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; TITLE OF INVENTION: Genes identified as required for  
; TITLE OF INVENTION: proliferation of E. coli  
; FILE REFERENCE: ELITRA.009A  
; CURRENT APPLICATION NUMBER: US/09/741.669  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: US 60/173005  
; PRIOR FILING DATE: 1999-12-23  
; NUMBER OF SEQ ID NOS: 481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 251  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-09-741-669-45

Alignment Scores: Length: 251  
Pred. No.: 2.6e+03  
Score: 22.00  
Percent Similarity: 80.00%  
Best Local Similarity: 80.00%  
Query Match: 91.67%  
DB: 9

US-10-030-194A-6 (1-6) x US-09-741-669-45 (1-251)

Oy 1 GlyTyr\*\*\*ValGlu 5

Db 242 GGTACACCGTCGAA 228

RESULT 50

US-10-282-122A-1445/c  
; Sequence 1445, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Hasebeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282.122A  
; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1445  
; LENGTH: 251  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-282-122A-1445

Alignment Scores: Length: 251  
Pred. No.: 2.6e+03  
Score: 22.00  
Percent Similarity: 80.00%  
Best Local Similarity: 80.00%  
Query Match: 91.67%  
DB: 16

US-10-030-194A-6 (1-6) x US-10-282-122A-1445 (1-251)

Oy 1 GlyTyr\*\*\*ValGlu 5

Db 242 GGTACACCGTCGAA 228

RESULT 51

US-10-424-599-85528  
; Sequence 85528, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424.599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 85528  
; LENGTH: 253  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_48246C.1  
US-10-424-599-85528

Alignment Scores: Length: 253  
Pred. No.: 2.63e+03  
Score: 22.00  
Percent Similarity: 80.00%  
Best Local Similarity: 80.00%  
Query Match: 91.67%  
DB: 16

```
US-10-030-194A-6 (1-6) x US-10-424-599-85528 (1-253)
QY 1 GlyTyr***ValGlu 5
DB 156 GGGTATTAGCGTGGAA 170

RESULT 52
US-10-437-963-86065/c
; Sequence 86065, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 86065
; LENGTH: 262
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85141C.1
US-10-437-963-86065

Alignment Scores: 2.73e+03 Length: 262
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-437-963-86065 (1-262)
QY 1 GlyTyr***ValGlu 5
DB 114 GGTATTACTCGTCGAG 100

RESULT 53
US-10-424-599-121254/c
; Sequence 121254, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 121254
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_8049C.1
US-10-424-599-121254

Alignment Scores: 2.75e+03 Length: 264
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 17 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-764-872-548 (1-265)
QY 1 GlyTyr***ValGlu 5
DB 52 GGTATTTCAGTGGAG 66

RESULT 55
US-09-764-872-37
; Sequence 37, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (236)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (267)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-872-37

Alignment Scores: 2.78e+03 Length: 267
Pred. No.: 22.00 Matches: 4
Score: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 10 Gaps: 0
```

Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 10 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-764-872-37 (1-267)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 55 GGTATTAGTGGAG 69

## RESULT 56

US-10-424-599-73754  
; Sequence 73754, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 73754  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37614C.1  
US-10-424-599-73754

Alignment Scores:  
Pred. No.: 2.78e+03 Length: 267  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-73754 (1-267)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 172 GGTATTAGTGGAA 186

## RESULT 57

US-10-424-599-58619/c  
; Sequence 58619, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 58619  
; LENGTH: 271  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_23945C.1  
US-10-424-599-58619

Alignment Scores:

Pred. No.: 2.83e+03 Length: 271  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-58619 (1-271)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 100 GGATATAGTGTGAG 86

## RESULT 58

US-10-140-463-1  
; Sequence 1, Application US/10140463  
; Publication No. US20030036522A1  
; GENERAL INFORMATION:  
; APPLICANT: Price, Jack  
; APPLICANT: Uwanogho, Dafe  
; TITLE OF INVENTION: Identification of Cells for Transplantation  
; FILE REFERENCE: GJE-36XCI  
; CURRENT APPLICATION NUMBER: US/10/140,463  
; CURRENT FILING DATE: 2002-05-06  
; PRIOR APPLICATION NUMBER: US/09/696,569  
; PRIOR FILING DATE: 2000-10-25  
; PRIOR APPLICATION NUMBER: 9925210.8  
; PRIOR FILING DATE: 1999-10-25  
; PRIOR APPLICATION NUMBER: 60/170,692  
; PRIOR FILING DATE: 1999-12-14  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 273  
; TYPE: DNA  
; ORGANISM: Mouse  
US-10-140-463-1

Alignment Scores:  
Pred. No.: 2.85e+03 Length: 273  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 14 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-140-463-1 (1-273)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 126 GGATATGCCGTGAA 140

## RESULT 59

US-10-424-599-78977  
; Sequence 78977, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 78977  
; LENGTH: 278  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure

```

; LOCATION: (1)..(278)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_WRT3847_42331C.1
US-10-424-599-78977

Alignment Scores:
Pred. No.: 2.9e+03 278
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-78977 (1-278)

Qy 1 GlyTyr***ValGlu 5
Db 140 GGTACACAGTGGAG 154

RESULT 60
US-09-741-669-46/c
; Sequence 46, Application US/09741669
; Patent No. US2002022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; TITLE OF INVENTION: Genes identified as required for
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-741-669-46

Alignment Scores:
Pred. No.: 2.99e+03 286
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 9 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-741-669-46 (1-286)

Qy 1 GlyTyr***ValGlu 5
Db 222 GGTACACCGTCGAA 208

RESULT 61
US-10-282-122A-1421/c
; Sequence 1421, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

```

```

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1421
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-282-122A-1421

Alignment Scores:
Pred. No.: 2.99e+03 286
Score: 22.00 Matches: 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-282-122A-1421 (1-286)

Qy 1 GlyTyr***ValGlu 5
Db 222 GGTACACCGTCGAA 208

RESULT 62
US-10-242-535A-46051
; Sequence 46051, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 46051
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Human

```



```
US-10-242-535A-46051
Alignment Scores:
Pred. No.:      3.03e+03      Length:      289
Score:          22.00         Matches:      4
Percent Similarity: 80.00%    Conservaive: 0
Best Local Similarity: 80.00% Mismatches:  1
Query Match:    91.67%       Indels:      0
DB:            16           Gaps:      0

US-10-030-194A-6 (1-6) x US-10-242-535A-46051 (1-289)

Qy      1 GlyTyr***ValGlu 5
Db      35 GGATATGCTGTAGAG 49

RESULT 63
US-10-424-599-29094/c
; Sequence 29094, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Zhou Yihua
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 29094
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_126272C.1
US-10-424-599-29094

Alignment Scores:
Pred. No.:      3.03e+03      Length:      289
Score:          22.00         Matches:      4
Percent Similarity: 80.00%    Conservaive: 0
Best Local Similarity: 80.00% Mismatches:  1
Query Match:    91.67%       Indels:      0
DB:            16           Gaps:      0

US-10-030-194A-6 (1-6) x US-10-424-599-29094 (1-289)

Qy      1 GlyTyr***ValGlu 5
Db      173 GGTTATAGTGTGAG 159

RESULT 64
US-10-085-783A-46051
; Sequence 46051, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 46051
```

```
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-46051

Alignment Scores:
Pred. No.:      3.03e+03      Length:      289
Score:          22.00         Matches:      4
Percent Similarity: 80.00%    Conservaive: 0
Best Local Similarity: 80.00% Mismatches:  1
Query Match:    91.67%       Indels:      0
DB:            16           Gaps:      0

US-10-030-194A-6 (1-6) x US-10-085-783A-46051 (1-289)

Qy      1 GlyTyr***ValGlu 5
Db      35 GGATATGCTGTAGAG 49

RESULT 65
US-10-424-599-18133
; Sequence 18133, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 18133
; LENGTH: 298
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_116378C.1
US-10-424-599-18133

Alignment Scores:
Pred. No.:      3.13e+03      Length:      298
Score:          22.00         Matches:      4
Percent Similarity: 80.00%    Conservaive: 0
Best Local Similarity: 80.00% Mismatches:  1
Query Match:    91.67%       Indels:      0
DB:            16           Gaps:      0

US-10-030-194A-6 (1-6) x US-10-424-599-18133 (1-298)

Qy      1 GlyTyr***ValGlu 5
Db      182 GGGTATGCTGTGGAG 196

RESULT 66
US-10-076-555-87
; Sequence 87, Application US/10076555
; Publication No. US20030065156A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominiguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
```

APPLICANT: Drmanac, Radoje  
APPLICANT: Crkvenjakov, Radomir  
APPLICANT: Dickson, Mark  
APPLICANT: Drmanac, Snezana  
APPLICANT: Labat, Ivan  
APPLICANT: Leshkowitz, Dena  
APPLICANT: Kita, David  
APPLICANT: Garcia, Veronica  
APPLICANT: Jones, Lee William  
APPLICANT: Stache-Crain, Birgit  
TITLE OF INVENTION: Novel Human Genes and Gene Expression  
FILE OF INVENTION: Products I  
FILE REFERENCE: 2300-1480  
CURRENT APPLICATION NUMBER: US/10/076,555  
CURRENT FILING DATE: 2002-02-15  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/217,471  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-21  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/068,755  
PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-23  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/080,664  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-03  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/105,234  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-21  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/105,877  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-27  
NUMBER OF SEQ ID NOS: 844  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 87  
LENGTH: 300  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-076-555-87

Alignment Scores:  
Pred. No.: 3.15e+03 Length: 300  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 14 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-076-555-87 (1-300)

Qy 1 GlyTyr\*\*\*Valglu 5  
Db 214 GGTATGCTGTGGAG 228

RESULT 67  
US-09-864-408A-1975/c  
Sequence 1975, Application US/09864408A  
Publication No. US20040009474A1  
GENERAL INFORMATION:  
APPLICANT: Leach, Martin D.  
TITLE OF INVENTION: Shinkets, Richard A.  
FILE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encoded  
FILE REFERENCE: 21402-012  
CURRENT APPLICATION NUMBER: US/09/864,408A  
CURRENT FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 60/206,690  
PRIOR FILING DATE: 2000-05-24  
NUMBER OF SEQ ID NOS: 9068  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1975  
LENGTH: 309  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-864-408A-1975

Alignment Scores:  
Pred. No.: 3.25e+03 Length: 309  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1

Query Match: 91.67% Indels: 0  
DB: 11 Gaps: 0  
US-10-030-194A-6 (1-6) x US-09-864-408A-1975 (1-309)

Qy 1 GlyTyr\*\*\*Valglu 5  
Db 119 GGATATTCTGTGGA 105

RESULT 68  
US-10-424-599-104737/c  
Sequence 104737, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J  
APPLICANT: Kovalic, David K  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 104737  
LENGTH: 310  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_65593C.1  
US-10-424-599-104737

Alignment Scores:  
Pred. No.: 3.26e+03 Length: 310  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-104737 (1-310)

Qy 1 GlyTyr\*\*\*Valglu 5  
Db 138 GGTACACAGTGGAG 124

RESULT 69  
US-10-437-963-28622  
Sequence 28622, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 28622  
LENGTH: 316  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_33202C.1  
US-10-437-963-28622

Alignment Scores:  
Pred. No.: 3.33e+03 Length: 316  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 17 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-437-963-28622 (1-316)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 103 CGATACGCGATCGAG 117

## RESULT 70

US-09-864-408A-3483  
; Sequence 3483, Application US/09864408A  
; Publication No. US20040009474A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Shimkets, Richard A.

; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encoded  
; FILE REFERENCE: 21402-012  
; CURRENT APPLICATION NUMBER: US/09/864, 408A  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/206,690  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 9068  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3483  
; LENGTH: 318  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-864-408A-3483

Alignment Scores:  
Pred. No.: 3.35e+03 Length: 318  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 11 Gaps: 0

US-10-030-194A-6 (1-6) x US-09-864-408A-3483 (1-318)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 165 CGATATTCTGTGAA 179

## RESULT 71

US-10-424-599-73534/c  
; Sequence 73534, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 73534  
; LENGTH: 318  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37416C.1  
US-10-424-599-73534

Alignment Scores:

Pred. No.: 3.35e+03 Length: 318  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-73534 (1-318)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 178 GGGTACTCCGTGAA 164

## RESULT 72

US-10-424-599-102597/c  
; Sequence 102597, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 102597  
; LENGTH: 319  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_63663C.1  
US-10-424-599-102597

Alignment Scores:  
Pred. No.: 3.36e+03 Length: 319  
Score: 22.00 Matches: 4  
Percent Similarity: 80.00% Conservative: 0  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 91.67% Indels: 0  
DB: 16 Gaps: 0

US-10-030-194A-6 (1-6) x US-10-424-599-102597 (1-319)

Qy 1 GlyTyr\*\*\*ValGlu 5  
|||||  
Db 171 GGGTATTCTGTGAA 157

## RESULT 73

US-10-011-095-9  
; Sequence 9, Application US/10011095  
; Publication No. US20030045682A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Hubert, Rene S.

; APPLICANT: Leong, Kahan  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Saffran, Douglas C.  
; APPLICANT: Mitchell, Steve Chappell  
; TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)  
; FILE REFERENCE: 511582001610  
; CURRENT APPLICATION NUMBER: US/10/011,095  
; CURRENT FILING DATE: 2001-12-06  
; PRIOR APPLICATION NUMBER: 09/323,873  
; PRIOR FILING DATE: 1999-06-01  
; PRIOR APPLICATION NUMBER: 60/087,520  
; PRIOR FILING DATE: 1998-06-01  
; PRIOR APPLICATION NUMBER: 60/091,183  
; PRIOR FILING DATE: 1998-06-30  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 9
; LENGTH: 322
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-011-095-9

Alignment Scores:
Pred. No.: 3.4e+03 322
Score: 22.00 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
Gaps: 14 0

US-10-030-194A-6 (1-6) x US-10-011-095-9 (1-322)

Qy 1 GlyTyr***ValGlu 5
Db 224 GGTACAGTGTAGAA 238

RESULT 74
US-10-010-667A-9
; Sequence 9, Application US/10010667A
; Publication No. US2003005217A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001601
; CURRENT APPLICATION NUMBER: US/10/010,667A
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 322
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-010-667A-9

Alignment Scores:
Pred. No.: 3.4e+03 322
Score: 22.00 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
Gaps: 14 0

US-10-030-194A-6 (1-6) x US-10-010-667A-9 (1-322)

Qy 1 GlyTyr***ValGlu 5
Db 224 GGTACAGTGTAGAA 238

RESULT 75
US-10-165-044-11
; Sequence 11, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran

```

```

; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappell Mitchell
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 51158-20016.02
; CURRENT APPLICATION NUMBER: US/10/165,044
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 322
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-165-044-11

Alignment Scores:
Pred. No.: 3.4e+03 322
Score: 22.00 4
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 91.67% Indels: 0
Gaps: 15 0

US-10-030-194A-6 (1-6) x US-10-165-044-11 (1-322)

Qy 1 GlyTyr***ValGlu 5
Db 224 GGTACAGTGTAGAA 238

Search completed: November 3, 2004, 19:57:16
Job time : 292 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:14:26 ; Search time 113 Seconds  
(without alignments)  
19.048 Million cell updates/sec

Title: US-10-030-194A-6

Perfect score: 24

Sequence: 1 GYVEX 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : A\_Geneseq\_23Sep04.\*

1: geneseqp1980s.\*

2: geneseqp1990s.\*

3: geneseqp2000s.\*

4: geneseqp2001s.\*

5: geneseqp2002s.\*

6: geneseqp2003as.\*

7: geneseqp2003bs.\*

8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	91.7	20	2	Aar15552 Immunopep
2	22	91.7	45	7	Adf70007 AcmaA-type
3	22	91.7	46	2	Aaw75574 M. tuberc
4	22	91.7	46	2	Aaw75573 M. tuberc
5	22	91.7	47	2	Aar92887 Mycobacte
6	22	91.7	47	6	Abg74428 M. tuberc
7	22	91.7	68	2	Aar63236 Neural th
8	22	91.7	68	2	Aar63236 Neural th
9	22	91.7	68	5	Aae29149 Neural th
10	22	91.7	68	6	Abg63247 68 amino
11	22	91.7	68	6	Abu02979 Human neu
12	22	91.7	68	6	Abp59229 Human 68
13	22	91.7	68	6	Aae33196 Neural th
14	22	91.7	68	6	Abj19452 68-mer ne
15	22	91.7	68	6	Abd37526 Neural th
16	22	91.7	68	7	Adl96027 Human neu
17	22	91.7	78	2	Aaw02576 E. coli p
18	22	91.7	78	2	Aar95919 E. coli p
19	22	91.7	83	2	Aay60563 Human nor
20	22	91.7	104	2	Aar91284 LERK-6 ex
21	22	91.7	104	2	Aay06821 Amino aci
22	22	91.7	113	7	Adh66927 Enterococ
23	22	91.7	125	3	Aag30135 Arabidops
24	22	91.7	140	7	Adh86269 Enterococ
25	22	91.7	143	7	Ab084239 Pseudomon

26	22	91.7	147	5	AAE15402
27	22	91.7	148	5	ABB48719
28	22	91.7	148	6	ABU32695
29	22	91.7	153	4	AAB99858
30	22	91.7	153	5	ABP26146
31	22	91.7	161	7	ABO76341
32	22	91.7	164	7	ADH86858
33	22	91.7	173	4	AAU55493
34	22	91.7	173	6	ABMS2012
35	22	91.7	179	6	ABU48693
36	22	91.7	180	3	AAG30134
37	22	91.7	184	2	AAW02587
38	22	91.7	184	2	AAAR91283
39	22	91.7	184	2	AAAY06820
40	22	91.7	199	8	ADO06904
41	22	91.7	200	4	ABG01444
42	22	91.7	201	4	AAAB74333
43	22	91.7	201	4	AAAG93162
44	22	91.7	203	7	ABO69448
45	22	91.7	205	7	ABO71975
46	22	91.7	206	4	AAAB79546
47	22	91.7	209	2	AAAR94766
48	22	91.7	209	2	AAAW71006
49	22	91.7	213	2	AAAY06822
50	22	91.7	224	5	AAO14749
51	22	91.7	232	7	ADC96464
52	22	91.7	235	6	ABM68208
53	22	91.7	238	5	ABG91554
54	22	91.7	238	7	ADM25747
55	22	91.7	246	2	AAW99400
56	22	91.7	250	6	ADA34441
57	22	91.7	264	4	ABB61132
58	22	91.7	270	7	AAO24038
59	22	91.7	270	7	AAO24039
60	22	91.7	277	2	AAW18875
61	22	91.7	282	2	AAV13560
62	22	91.7	283	5	ABP66166
63	22	91.7	296	7	ADF41714
64	22	91.7	297	2	AAW76698
65	22	91.7	297	2	AAW33009
66	22	91.7	297	2	AAW64449
67	22	91.7	297	2	AAW73004
68	22	91.7	297	6	AAE34500
69	22	91.7	297	8	ADF82797
70	22	91.7	297	8	ADN59253
71	22	91.7	298	4	ABG23387
72	22	91.7	300	5	ABP38577
73	22	91.7	302	5	ABB07518
74	22	91.7	307	4	AAAG82514
75	22	91.7	307	6	ABU43080
76	22	91.7	308	6	ABU22922
77	22	91.7	308	6	ADA34825
78	22	91.7	310	6	ABR38908
79	22	91.7	313	7	ABO84090
80	22	91.7	315	4	AAAG90121
81	22	91.7	321	5	ABG97426
82	22	91.7	322	6	ABU43449
83	22	91.7	322	6	ABU43410
84	22	91.7	323	4	AAAM00957
85	22	91.7	325	4	AAU34033
86	22	91.7	325	4	AAU36761
87	22	91.7	325	4	AAAG82118
88	22	91.7	325	4	AAAG81802
89	22	91.7	325	6	ABR64015
90	22	91.7	325	6	ABR64014
91	22	91.7	325	6	ABU42765
92	22	91.7	325	6	ABU15980
93	22	91.7	325	6	ABM73273
94	22	91.7	333	6	ABP80856
95	22	91.7	338	7	ABO63655
96	22	91.7	341	4	ABB65835
97	22	91.7	344	4	AAU33563
98	22	91.7	344	6	AAU15587

Aae15402	HPV 16 E2
Abb48719	Listeria
Abu32695	Protein e
Aab99858	Physcomit
Abp26146	Streptoco
Ab076341	Pseudomon
Adh86858	Enterococ
Aau55493	Propionib
Abms2012	Propionib
Abu48693	Protein e
Aag30134	Arabidops
Aaw02587	Lerk-6 pr
Aar91283	LERK-6 pr
Aay06820	Murine LE
Ado06904	Lactate d
Abg01444	Novel hum
Ab74333	Human pap
Aag93162	C glutami
Ab069448	Pseudomon
Ab071975	Pseudomon
Aab79546	Corynebac
Aar94766	Mouse EPH
Aaw71006	Amino aci
Aay06822	Human LER
Aa014749	Corynebac
Adc96464	E. faectiu
Abm68208	Photorhab
Abg91554	Purine/py
Adm25747	Hyperther
Aaw99400	S. antibio
Ada34441	Acinetoba
Abb61132	Drosophil
Aao24038	Aspergill
Aao24039	Aspergill
Aaw18875	Babesia e
Aav13560	S. haemol
Abp66166	Bifidobac
Adf41714	Bacillus
Aaw76698	Lipase 3
Aaw33009	Aspergill
Aaw64449	A. foetid
Aaw73004	Aspergill
Aae34500	Aspergill
Adf82797	Aspergill
Adn59253	Aspergill
Abg23387	Novel hum
Abp38577	Staphyloc
Abb07518	Human dru
Aag82514	S. epider
Abu43080	Protein e
Abu22922	Protein e
Ada34825	Acinetoba
Ab38908	pBAL ORF
Ab084090	C glutami
Aag90121	C glutami
Abg97426	M. echino
Abu43449	Protein e
Abu43410	Protein e
Aam00957	Human bon
Aau34033	Staphyloc
Aau36761	Staphyloc
Aag82118	S. epider
Aag81802	S. epider
Ab64015	Experiment
Ab64014	Predicted
Abu42765	Protein e
Abu15980	Protein e
Abm73273	Staphyloc
Abp80856	N. gonorr
Ab063655	Klebsiell
Abb65835	Drosophil
Aau33563	Pseudomon
Abu15587	Protein e

99	22	91.7	345	5	ABP38929	Staphyloc	172	22	91.7	460	4	AAM40318	Human pol
100	22	91.7	347	6	AAU58802	Propionib	173	22	91.7	460	5	ABBS4552	Lactococc
101	22	91.7	347	6	ARM55321	Propionib	174	22	91.7	461	3	AAG14437	Arabidops
102	22	91.7	349	7	ABO78798	Pseudomon	175	22	91.7	462	4	ABAB74506	White lup
103	22	91.7	352	3	AAB44828	Gene 49 h	176	22	91.7	462	4	AAM41054	Human pol
104	22	91.7	352	7	ADM26449	Hyperther	177	22	91.7	462	5	ABB80757	L. albus
105	22	91.7	354	7	ADC95552	E. faeciu	178	22	91.7	462	5	ABB80757	L. albus
106	22	91.7	355	6	ABU34791	Protein e	179	22	91.7	462	6	ABU48172	Listeria
107	22	91.7	360	5	ABP26427	Streptoco	180	22	91.7	465	6	ABU48172	Listeria
108	22	91.7	360	5	ABU45573	Protein e	181	22	91.7	465	8	AAM43556	Human pol
109	22	91.7	363	5	ABB47702	Listeria	182	22	91.7	466	2	ADM24577	Human pol
110	22	91.7	363	6	ABU32432	Protein e	183	22	91.7	466	2	AY05882	Potato ph
111	22	91.7	365	2	AAW12095	Human pap	184	22	91.7	466	3	ABBS4436	Arabidops
112	22	91.7	365	2	AAW50005	Human pap	185	22	91.7	466	3	ABBS4436	Arabidops
113	22	91.7	365	4	ABN98418	Human pap	186	22	91.7	468	5	ABG29673	Arabidops
114	22	91.7	365	4	AAB35191	Human pap	187	22	91.7	468	5	ABB91771	Herbicida
115	22	91.7	365	4	AAB35192	Human pap	188	22	91.7	469	6	ABP78967	N. gonorr
116	22	91.7	365	5	AAU10809	Human pap	189	22	91.7	469	6	ABU73362	Protein e
117	22	91.7	365	5	ABG70653	HPV-16/HP	190	22	91.7	477	6	ABU28782	Protein e
118	22	91.7	365	5	ABG70651	Human pap	191	22	91.7	486	4	AAU19923	Arabidops
119	22	91.7	365	6	ABO07151	Novel hum	192	22	91.7	488	4	ABG29672	Arabidops
120	22	91.7	365	6	ABO07150	Novel hum	193	22	91.7	488	4	ABG29672	Arabidops
121	22	91.7	365	6	AAO16779	Mutant HP	194	22	91.7	488	4	ABG29672	Arabidops
122	22	91.7	365	6	AAO16776	Mutant HP	195	22	91.7	488	4	ABG29672	Arabidops
123	22	91.7	365	6	AAO16777	Mutant HP	196	22	91.7	488	4	ABG29672	Arabidops
124	22	91.7	365	6	AAO16778	Mutant HP	197	22	91.7	488	4	ABG29672	Arabidops
125	22	91.7	365	6	AAO16780	Human pap	198	22	91.7	488	4	ABG29672	Arabidops
126	22	91.7	365	6	ABP98182	Amino aci	199	22	91.7	523	8	ADM99121	Bacterial
127	22	91.7	365	6	ADC64522	HPV 16 tr	200	22	91.7	523	8	ADM99121	Bacterial
128	22	91.7	365	7	ADC64525	HPV 16 tr	201	22	91.7	523	8	ADM99121	Bacterial
129	22	91.7	365	7	ADC64522	HPV 16 tr	202	22	91.7	523	8	ADM99121	Bacterial
130	22	91.7	372	4	AAAB98432	Human pap	203	22	91.7	530	7	ADBS8009	Human PRO
131	22	91.7	373	6	ABP79526	N. gonorr	204	22	91.7	530	7	ADBS8009	Human PRO
132	22	91.7	374	6	ABM64688	Propionib	205	22	91.7	530	7	ADBS8009	Human PRO
133	22	91.7	375	4	ABG17810	Novel hum	206	22	91.7	530	7	ADBS8009	Human PRO
134	22	91.7	375	4	ABG18221	Novel hum	207	22	91.7	530	7	ADBS8009	Human PRO
135	22	91.7	375	8	ADN60451	B. lichen	208	22	91.7	530	7	ADBS8009	Human PRO
136	22	91.7	377	2	AAW76330	Interleuk	209	22	91.7	530	7	ADBS8009	Human PRO
137	22	91.7	377	2	AAW76330	Interleuk	210	22	91.7	530	7	ADBS8009	Human PRO
138	22	91.7	379	4	ABE63073	Drosophil	211	22	91.7	530	7	ADBS8009	Human PRO
139	22	91.7	379	8	ADO80285	Neisseria	212	22	91.7	530	7	ADBS8009	Human PRO
140	22	91.7	380	5	ABB90775	Human Tum	213	22	91.7	530	7	ADBS8009	Human PRO
141	22	91.7	380	6	ABU54482	Human nor	214	22	91.7	530	7	ADBS8009	Human PRO
142	22	91.7	381	4	ABU79552	Corynebac	215	22	91.7	530	7	ADBS8009	Human PRO
143	22	91.7	382	5	ABG97431	S. Ghanae	216	22	91.7	530	7	ADBS8009	Human PRO
144	22	91.7	382	5	ABJ69625	Human hea	217	22	91.7	530	7	ADBS8009	Human PRO
145	22	91.7	382	7	ABO81133	Pseudomon	218	22	91.7	530	7	ADBS8009	Human PRO
146	22	91.7	383	4	AAG90479	C. glutami	219	22	91.7	530	7	ADBS8009	Human PRO
147	22	91.7	385	7	ADM25471	Hyperther	220	22	91.7	530	7	ADBS8009	Human PRO
148	22	91.7	389	1	AAFP70668	D-alanine	221	22	91.7	530	7	ADBS8009	Human PRO
149	22	91.7	389	2	AAV13558	B. subtil	222	22	91.7	530	7	ADBS8009	Human PRO
150	22	91.7	390	7	ABO75771	Pseudomon	223	22	91.7	530	7	ADBS8009	Human PRO
151	22	91.7	394	4	AAG90519	C. glutami	224	22	91.7	530	7	ADBS8009	Human PRO
152	22	91.7	396	4	AAG80925	Lipoic ac	225	22	91.7	530	7	ADBS8009	Human PRO
153	22	91.7	403	7	ADE60770	Human pro	226	22	91.7	530	7	ADBS8009	Human PRO
154	22	91.7	403	8	ADL70341	Crenarcha	227	22	91.7	530	7	ADBS8009	Human PRO
155	22	91.7	404	4	AAW39268	Human pol	228	22	91.7	530	7	ADBS8009	Human PRO
156	22	91.7	404	7	ABU62784	Human pro	229	22	91.7	530	7	ADBS8009	Human PRO
157	22	91.7	404	7	ABG75061	Human pro	230	22	91.7	530	7	ADBS8009	Human PRO
158	22	91.7	405	3	AAB26155	B. cereus	231	22	91.7	530	7	ADBS8009	Human PRO
159	22	91.7	406	4	ABB66742	Drosophil	232	22	91.7	530	7	ADBS8009	Human PRO
160	22	91.7	413	5	ABP26813	Streptoco	233	22	91.7	530	7	ADBS8009	Human PRO
161	22	91.7	421	5	ADJ181630	Transloca	234	22	91.7	530	7	ADBS8009	Human PRO
162	22	91.7	423	5	ABP65596	Bifidobac	235	22	91.7	530	7	ADBS8009	Human PRO
163	22	91.7	424	6	ABP57621	S. muraya	236	22	91.7	530	7	ADBS8009	Human PRO
164	22	91.7	426	3	ABO67185	Klebsiell	237	22	91.7	530	7	ADBS8009	Human PRO
165	22	91.7	431	7	ABO67185	Klebsiell	238	22	91.7	530	7	ADBS8009	Human PRO
166	22	91.7	434	4	AAW95565	Human pro	239	22	91.7	530	7	ADBS8009	Human PRO
167	22	91.7	439	6	ABU27953	Protein e	240	22	91.7	530	7	ADBS8009	Human PRO
168	22	91.7	442	8	ADJ49617	Oil-assoc	241	22	91.7	530	7	ADBS8009	Human PRO
169	22	91.7	444	8	ADM47748	Thermococ	242	22	91.7	530	7	ADBS8009	Human PRO
170	22	91.7	446	5	ABP29219	Streptoco	243	22	91.7	530	7	ADBS8009	Human PRO
171	22	91.7	460	2	AAR54216	L.lactis	244	22	91.7	530	7	ADBS8009	Human PRO

245	22	91.7	541	6	ABU88728	Novel hum	318	22	91.7	541	6	ABO09627	Human sec
246	22	91.7	541	6	ABU83423	Human sec	319	22	91.7	541	6	ABO10899	Human sec
247	22	91.7	541	6	ABO06224	Novel hum	320	22	91.7	541	6	ABO10899	Human sec
248	22	91.7	541	6	ABR59260	Human sec	321	22	91.7	541	6	ABU87561	Human PRO
249	22	91.7	541	6	ABR59260	Human sec	322	22	91.7	541	6	ABU91429	Human PRO
250	22	91.7	541	6	ABO09322	Human sec	323	22	91.7	541	6	ABU84643	Human sec
251	22	91.7	541	6	ABO19186	Novel hum	324	22	91.7	541	6	ABR69733	Human sec
252	22	91.7	541	6	ABO11204	Human sec	325	22	91.7	541	6	ABU80110	Human PRO
253	22	91.7	541	6	ABR66822	Human sec	326	22	91.7	541	6	ABU92153	Novel hum
254	22	91.7	541	6	ABO16035	Human sec	327	22	91.7	541	6	ABU93379	Human PRO
255	22	91.7	541	6	ABO13741	Human sec	328	22	91.7	541	6	ABO09932	Human sec
256	22	91.7	541	6	ABU65644	Human sec	329	22	91.7	541	6	ABO09017	Human sec
257	22	91.7	541	6	ABO07492	Human PRO	330	22	91.7	541	6	ABU10859	Human PRO
258	22	91.7	541	6	ABO03679	Human sec	331	22	91.7	541	6	ABU10585	Human sec
259	22	91.7	541	6	ABR67127	Human sec	332	22	91.7	541	6	ABU81611	Novel hum
260	22	91.7	541	6	ABO15730	Human sec	333	22	91.7	541	6	ABU95594	Human PRO
261	22	91.7	541	6	ABU56011	Human sec	334	22	91.7	541	6	ABU96803	Novel hum
262	22	91.7	541	6	ABU65339	Human PRO	335	22	91.7	541	6	ABR70648	Human sec
263	22	91.7	541	6	ABU95284	Novel hum	336	22	91.7	541	6	ABO04999	Novel hum
264	22	91.7	541	6	ABU71187	Human PRO	337	22	91.7	541	6	ABO08407	Human sec
265	22	91.7	541	6	ABO07797	Human PRO	338	22	91.7	541	6	ABU88550	Human sec
266	22	91.7	541	6	ABR70038	Human sec	339	22	91.7	541	6	ABO34064	Human PRO
267	22	91.7	541	6	ABR69371	Human sec	340	22	91.7	541	6	ABO05614	Human sec
268	22	91.7	541	6	ABO01512	Human PRO	341	22	91.7	541	6	ABR74003	Human sec
269	22	91.7	541	6	ABU81314	Human PRO	342	22	91.7	541	6	ABR95595	Human sec
270	22	91.7	541	6	ABR60111	Human sec	343	22	91.7	541	6	ABR80892	Human sec
271	22	91.7	541	6	ABR67846	Human sec	344	22	91.7	541	6	ABR81197	Human sec
272	22	91.7	541	6	ABR65234	Human sec	345	22	91.7	541	6	ABM00893	Human sec
273	22	91.7	541	6	ABR68456	Human sec	346	22	91.7	541	6	ABR88495	Human sec
274	22	91.7	541	6	ABR71868	Human sec	347	22	91.7	541	6	ABM77316	Human sec
275	22	91.7	541	6	ABU59238	Human sec	348	22	91.7	541	6	ABO28800	Human sec
276	22	91.7	541	6	ABU85348	Human PRO	349	22	91.7	541	6	ABO31545	Human sec
277	22	91.7	541	6	ABU89038	Human sec	350	22	91.7	541	6	ABM07962	Human sec
278	22	91.7	541	6	ABU83118	Human sec	351	22	91.7	541	6	ABO40442	Human sec
279	22	91.7	541	6	ABU94974	Novel hum	352	22	91.7	541	6	ABO35867	Human PRO
280	22	91.7	541	6	ABU90522	Novel hum	353	22	91.7	541	6	ABO44006	Human PRO
281	22	91.7	541	6	ABU84033	Human sec	354	22	91.7	541	6	ADA77904	Human sec
282	22	91.7	541	6	ABU93684	Novel hum	355	22	91.7	541	6	ABM24801	Human sec
283	22	91.7	541	6	ABO25935	Human PRO	356	22	91.7	541	6	ABO03069	Human sec
284	22	91.7	541	6	ABR64929	Human sec	357	22	91.7	541	6	ABR90325	Human sec
285	22	91.7	541	6	ABR68761	Human sec	358	22	91.7	541	6	ABM17239	Human sec
286	22	91.7	541	6	ABO06577	Human sec	359	22	91.7	541	6	ABR94985	Human sec
287	22	91.7	541	6	ABR99122	Human sec	360	22	91.7	541	6	ABR95290	Human sec
288	22	91.7	541	6	ABU57006	Human PRO	361	22	91.7	541	6	ABO21528	Human sec
289	22	91.7	541	6	ABU85958	Novel hum	362	22	91.7	541	6	ABR97792	Human sec
290	22	91.7	541	6	ABU82245	Novel hum	363	22	91.7	541	6	ABR87580	Human sec
291	22	91.7	541	6	ABU87256	Human PRO	364	22	91.7	541	6	ABM77621	Human sec
292	22	91.7	541	6	ABU83728	Human sec	365	22	91.7	541	6	ABM27851	Human sec
293	22	91.7	541	6	ABO08102	Human PRO	366	22	91.7	541	6	ABM06132	Human sec
294	22	91.7	541	6	ABU81813	Novel hum	367	22	91.7	541	6	ABM03638	Human sec
295	22	91.7	541	6	ABU65977	Novel hum	368	22	91.7	541	6	ABM35089	Human sec
296	22	91.7	541	6	ABR59806	Human sec	369	22	91.7	541	6	ABM26326	Human sec
297	22	91.7	541	6	ABU93994	Novel hum	370	22	91.7	541	6	ABO48108	Human sec
298	22	91.7	541	6	ABU99847	Novel hum	371	22	91.7	541	6	ABR92850	Human sec
299	22	91.7	541	6	ABR65517	Human sec	372	22	91.7	541	6	ABO24611	Human sec
300	22	91.7	541	6	ABR90935	Human sec	373	22	91.7	541	6	ADA37692	Human sec
301	22	91.7	541	6	ABU58944	Human sec	374	22	91.7	541	6	ABM11622	Human sec
302	22	91.7	541	6	ABU94362	Human PRO	375	22	91.7	541	6	ABM02723	Human sec
303	22	91.7	541	6	ABU79244	Human PRO	376	22	91.7	541	6	ABM16019	Human sec
304	22	91.7	541	6	ABU86573	Human sec	377	22	91.7	541	6	ABM16019	Human sec
305	22	91.7	541	6	ABU86878	Novel hum	378	22	91.7	541	6	ABO27580	Human sec
306	22	91.7	541	6	ABU94667	Human PRO	379	22	91.7	541	6	ABM29071	Human sec
307	22	91.7	541	6	ABO04594	Human PRO	380	22	91.7	541	6	ABM07047	Human sec
308	22	91.7	541	6	ABR70343	Human sec	381	22	91.7	541	6	ABM21141	Human sec
309	22	91.7	541	6	ABU92322	Novel hum	382	22	91.7	541	6	ABM09487	Human sec
310	22	91.7	541	6	ABU98508	Human PRO	383	22	91.7	541	6	ABO41357	Human sec
311	22	91.7	541	6	ABR65907	Human sec	384	22	91.7	541	6	ABO36172	Human PRO
312	22	91.7	541	6	ABR64624	Human sec	385	22	91.7	541	6	ABO43701	Human PRO
313	22	91.7	541	6	ABU59387	Novel hum	386	22	91.7	541	6	ABM76097	Human sec
314	22	91.7	541	6	ABU79549	Human PRO	387	22	91.7	541	6	ABM25716	Human sec
315	22	91.7	541	6	ABU92940	Human sec	388	22	91.7	541	6	ABM26021	Human sec
316	22	91.7	541	6	ABU95899	Human PRO	389	22	91.7	541	6	ABM26021	Human sec
317	22	91.7	541	6	ABU91119	Novel hum	390	22	91.7	541	6	ADA21378	Human sec
	22	91.7	541	6	ABU90212	Novel hum	391	22	91.7	541	6	ABO03374	Human sec

391	22	91.7	541	6	ABO2459	Human sec	464	22	91.7	541	6	ABM75792	Human sec
392	22	91.7	541	6	ABR90630	Human sec	465	22	91.7	541	6	ABM34072	Human sec
393	22	91.7	541	6	ABR73698	Human sec	466	22	91.7	541	6	ABM34377	Human sec
394	22	91.7	541	6	ABO16950	Human sec	467	22	91.7	541	6	ABO20308	Human sec
395	22	91.7	541	6	ABR94375	Human sec	468	22	91.7	541	6	ABO21223	Human sec
396	22	91.7	541	6	ABR75882	Human sec	469	22	91.7	541	6	ABO22138	Human sec
397	22	91.7	541	6	ABR71258	Human sec	470	22	91.7	541	6	ABR96572	Human sec
398	22	91.7	541	6	ABR93155	Human sec	471	22	91.7	541	6	ADA94397	Human sec
399	22	91.7	541	6	ABR93460	Human sec	472	22	91.7	541	6	ABR85750	Human sec
400	22	91.7	541	6	ADA10165	Human sec	473	22	91.7	541	6	ABR99732	Human sec
401	22	91.7	541	6	ABR87885	Human sec	474	22	91.7	541	6	ABM00283	Human sec
402	22	91.7	541	6	ABO27885	Human sec	475	22	91.7	541	6	ABM00588	Human sec
403	22	91.7	541	6	ABO30020	Human PRO	476	22	91.7	541	6	ABO29715	Human sec
404	22	91.7	541	6	ABO33229	Human sec	477	22	91.7	541	6	ABM29376	Human sec
405	22	91.7	541	6	ABM04917	Human sec	478	22	91.7	541	6	ABO38307	Human sec
406	22	91.7	541	6	ABM08877	Human sec	479	22	91.7	541	6	ABO45607	Human PRO
407	22	91.7	541	6	ABO36477	Human sec	480	22	91.7	541	6	ABM20531	Human sec
408	22	91.7	541	6	ABO35562	Human sec	481	22	91.7	541	6	ADA81423	Human sec
409	22	91.7	541	6	ABO39527	Human sec	482	22	91.7	541	6	ABO16645	Human sec
410	22	91.7	541	6	ABM10402	Human sec	483	22	91.7	541	6	ABO18271	Human PRO
411	22	91.7	541	6	ABM11927	Human PRO	484	22	91.7	541	6	ABO22698	Human PRO
412	22	91.7	541	6	ABO52073	Human PRO	485	22	91.7	541	6	ABO23003	Human PRO
413	22	91.7	541	6	ABO52378	Human sec	486	22	91.7	541	6	ABR92545	Human sec
414	22	91.7	541	6	ABO23696	Human sec	487	22	91.7	541	6	ABR81502	Human sec
415	22	91.7	541	6	ADA17709	Human PRO	488	22	91.7	541	6	ABM77926	Human sec
416	22	91.7	541	6	ABR97182	Human sec	489	22	91.7	541	6	ABR89715	Human sec
417	22	91.7	541	6	ABR86970	Human sec	490	22	91.7	541	6	ABM26631	Human sec
418	22	91.7	541	6	ABM11012	Human sec	491	22	91.7	541	6	ABM13757	Human sec
419	22	91.7	541	6	ABM28156	Human sec	492	22	91.7	541	6	ABM13757	Human sec
420	22	91.7	541	6	ABO32155	Human sec	493	22	91.7	541	6	ABO28495	Human sec
421	22	91.7	541	6	ABM15282	Human sec	494	22	91.7	541	6	ABO30325	Human sec
422	22	91.7	541	6	ABM06437	Human sec	495	22	91.7	541	6	ABM07352	Human sec
423	22	91.7	541	6	ABM04248	Human sec	496	22	91.7	541	6	ABO37087	Human sec
424	22	91.7	541	6	ABM22361	Human sec	497	22	91.7	541	6	ABO41662	Human PRO
425	22	91.7	541	6	ABO07657	Human sec	498	22	91.7	541	6	ABO35257	Human sec
426	22	91.7	541	6	ABO40747	Human sec	499	22	91.7	541	6	ABM25106	Human sec
427	22	91.7	541	6	ABM35394	Human sec	500	22	91.7	541	6	ABO47498	Human sec
428	22	91.7	541	6	ABM33157	Human sec	501	22	91.7	541	6	ABO47803	Human sec
429	22	91.7	541	6	ABO52683	Human PRO	502	22	91.7	541	6	ABO48413	Human sec
430	22	91.7	541	6	ABO50243	Human sec	503	22	91.7	541	6	ABO51463	Human PRO
431	22	91.7	541	6	ABU99237	Human sec	504	22	91.7	541	6	ABO51768	Human PRO
432	22	91.7	541	6	ABO04289	Human sec	505	22	91.7	541	6	ABO50548	Human sec
433	22	91.7	541	6	ABM18459	Human sec	506	22	91.7	541	6	ABR79672	Human sec
434	22	91.7	541	6	ADA27817	Human sec	507	22	91.7	541	6	ABM16934	Human sec
435	22	91.7	541	6	ABR97487	Human sec	508	22	91.7	541	6	ABO17966	Human sec
436	22	91.7	541	6	ABR80587	Human sec	509	22	91.7	541	6	ABO20918	Human sec
437	22	91.7	541	6	ABR80587	Human sec	510	22	91.7	541	6	ABR96877	Human sec
438	22	91.7	541	6	ABM01198	Human sec	511	22	91.7	541	6	ADA38622	Human sec
439	22	91.7	541	6	ABR88800	Human sec	512	22	91.7	541	6	ABM12232	Human sec
440	22	91.7	541	6	ABM13452	Human sec	513	22	91.7	541	6	ABM16324	Human sec
441	22	91.7	541	6	ABM20836	Human sec	514	22	91.7	541	6	ABM16324	Human sec
442	22	91.7	541	6	ABO41967	Human sec	515	22	91.7	541	6	ABM24191	Human sec
443	22	91.7	541	6	ABO42577	Human sec	516	22	91.7	541	6	ABM14672	Human sec
444	22	91.7	541	6	ABM10097	Human sec	517	22	91.7	541	6	ABM04553	Human sec
445	22	91.7	541	6	ABO38612	Human sec	518	22	91.7	541	6	ABM06742	Human sec
446	22	91.7	541	6	ABM32852	Human sec	519	22	91.7	541	6	ABO39222	Human sec
447	22	91.7	541	6	ABM22666	Human sec	520	22	91.7	541	6	ABO39222	Human sec
448	22	91.7	541	6	ABM74877	Human sec	521	22	91.7	541	6	ABM75487	Human sec
449	22	91.7	541	6	ADAT9696	Human sec	522	22	91.7	541	6	ABM25411	Human sec
450	22	91.7	541	6	ABR96267	Human sec	523	22	91.7	541	6	ABM19921	Human PRO
451	22	91.7	541	6	ABM02418	Human sec	524	22	91.7	541	6	ABO46827	Human PRO
452	22	91.7	541	6	ABR86360	Human sec	525	22	91.7	541	6	ABO47132	Human PRO
453	22	91.7	541	6	ABR86665	Human sec	526	22	91.7	541	6	ADA83221	Human sec
454	22	91.7	541	6	ABM16629	Human sec	527	22	91.7	541	6	ABR71563	Human sec
455	22	91.7	541	6	ABM29681	Human sec	528	22	91.7	541	6	ABR72173	Human sec
456	22	91.7	541	6	ABO29105	Human sec	529	22	91.7	541	6	ABR98512	Human sec
457	22	91.7	541	6	ABM23886	Human sec	530	22	91.7	541	6	ABR98512	Human sec
458	22	91.7	541	6	ABM23276	Human sec	531	22	91.7	541	6	ABR84835	Human sec
459	22	91.7	541	6	ABM22056	Human sec	532	22	91.7	541	6	ABR73393	Human sec
460	22	91.7	541	6	ABO37697	Human sec	533	22	91.7	541	6	ABR76487	Human sec
461	22	91.7	541	6	ABM28461	Human sec	534	22	91.7	541	6	ABR73088	Human sec
462	22	91.7	541	6	ABM28766	Human sec	535	22	91.7	541	6	ABM18154	Human sec
463	22	91.7	541	6	ABM66410	Human sec	536	22	91.7	541	6	ABO20613	Human sec



537	22	91.7	541	6	ABO25356	Human	PRO	610	22	91.7	541	7	ABO24001	Human	sec
538	22	91.7	541	6	ABO25661	Human	PRO	611	22	91.7	541	7	ABR93765	Human	sec
539	22	91.7	541	6	ABR94070	Human	sec	612	22	91.7	541	7	ABM01808	Human	sec
540	22	91.7	541	6	ADA92743	Human	sec	613	22	91.7	541	7	ABM78231	Human	sec
541	22	91.7	541	6	ABR79977	Human	sec	614	22	91.7	541	7	ABR90020	Human	sec
542	22	91.7	541	6	ABM11317	Human	sec	615	22	91.7	541	7	ADA22304	Human	sec
543	22	91.7	541	6	ABM11317	Human	sec	616	22	91.7	541	7	ABM27546	Human	sec
544	22	91.7	541	6	ABO32924	Human	PRO	617	22	91.7	541	7	ABM13147	Human	sec
545	22	91.7	541	6	ABO30630	Human	sec	618	22	91.7	541	7	ABO31850	Human	sec
546	22	91.7	541	6	ABO30935	Human	sec	619	22	91.7	541	7	ABM14062	Human	sec
547	22	91.7	541	6	ABM27241	Human	sec	620	22	91.7	541	7	ABM08267	Human	sec
548	22	91.7	541	6	ABM29986	Human	sec	621	22	91.7	541	7	ABO40137	Human	sec
549	22	91.7	541	6	ABM05522	Human	sec	622	22	91.7	541	7	ABM74572	Human	sec
550	22	91.7	541	6	ABM15587	Human	sec	623	22	91.7	541	7	ABM33767	Human	sec
551	22	91.7	541	6	ABM08572	Human	sec	624	22	91.7	541	7	ABM20226	Human	sec
552	22	91.7	541	6	ABO42272	Human	sec	625	22	91.7	541	7	ABO48718	Human	sec
553	22	91.7	541	6	ABO38002	Human	sec	626	22	91.7	541	7	ABO22520	Human	sec
554	22	91.7	541	6	ABO45912	Human	PRO	627	22	91.7	541	7	ABR72783	Human	sec
555	22	91.7	541	6	ABM66715	Human	sec	628	22	91.7	541	7	ABO15425	Human	sec
556	22	91.7	541	6	ABM66715	Human	sec	629	22	91.7	541	7	ABR85140	Human	sec
557	22	91.7	541	6	ABD20264	Human	sec	630	22	91.7	541	7	ABO15120	Human	sec
558	22	91.7	541	6	ABM19616	Human	sec	631	22	91.7	541	7	ABO17255	Human	sec
559	22	91.7	541	6	ABO49328	Human	sec	632	22	91.7	541	7	ABM17544	Human	sec
560	22	91.7	541	6	ABO49633	Human	sec	633	22	91.7	541	7	ADA06470	Human	sec
561	22	91.7	541	6	ADA78516	Human	sec	634	22	91.7	541	7	ADA39163	Human	sec
562	22	91.7	541	6	ABR88190	Human	sec	635	22	91.7	541	7	ABR85445	Human	sec
563	22	91.7	541	6	ABM26936	Human	sec	636	22	91.7	541	7	ABM77011	Human	sec
564	22	91.7	541	6	ABM03333	Human	sec	637	22	91.7	541	7	ABO28190	Human	sec
565	22	91.7	541	6	ABO39832	Human	sec	638	22	91.7	541	7	ABM22971	Human	sec
566	22	91.7	541	6	ABO49938	Human	sec	639	22	91.7	541	7	ABM30291	Human	sec
567	22	91.7	541	6	ABO50853	Human	sec	640	22	91.7	541	7	ABM21751	Human	sec
568	22	91.7	541	6	ABO5309	Human	sec	641	22	91.7	541	7	ABM21446	Human	sec
569	22	91.7	541	6	ABR74613	Human	sec	642	22	91.7	541	7	ABM14977	Human	sec
570	22	91.7	541	6	ABR77092	Human	sec	643	22	91.7	541	7	ABO41052	Human	sec
571	22	91.7	541	6	ABM17849	Human	sec	644	22	91.7	541	7	ABO36782	Human	sec
572	22	91.7	541	6	ABR95900	Human	sec	645	22	91.7	541	7	ABO37392	Human	sec
573	22	91.7	541	6	ABO21833	Human	sec	646	22	91.7	541	7	ABM75182	Human	sec
574	22	91.7	541	6	ABO20003	Human	sec	647	22	91.7	541	7	ABM33462	Human	sec
575	22	91.7	541	6	ABO24306	Human	sec	648	22	91.7	541	7	ABO46217	Human	PRO
576	22	91.7	541	6	ABR86055	Human	sec	649	22	91.7	541	7	ADA82587	Human	sec
577	22	91.7	541	6	ABM10707	Human	sec	650	22	91.7	541	7	ADB96189	Human	PRO
578	22	91.7	541	6	ABM76706	Human	sec	651	22	91.7	541	7	ABM31816	Human	sec
579	22	91.7	541	6	ABR89410	Human	sec	652	22	91.7	541	7	ABM31206	Human	sec
580	22	91.7	541	6	ABM12537	Human	sec	653	22	91.7	541	7	ADB85895	Human	sec
581	22	91.7	541	6	ABM05827	Human	sec	654	22	91.7	541	7	ABM32121	Human	sec
582	22	91.7	541	6	ABO34952	Human	PRO	655	22	91.7	541	7	ABM32426	Human	sec
583	22	91.7	541	6	ABO3028	Human	sec	656	22	91.7	541	7	ABM31511	Human	sec
584	22	91.7	541	6	ABM19006	Human	sec	657	22	91.7	541	7	ABM30901	Human	sec
585	22	91.7	541	6	ABM19311	Human	sec	658	22	91.7	541	7	ADC57661	Human	PRO
586	22	91.7	541	6	ABO46522	Human	sec	659	22	91.7	541	7	ADC55025	Human	PRO
587	22	91.7	541	6	ABO49023	Human	sec	660	22	91.7	541	7	ADC11892	Human	PRO
588	22	91.7	541	6	ABR69066	Human	sec	661	22	91.7	541	7	ADC56314	Human	PRO
589	22	91.7	541	6	ABR89105	Human	sec	662	22	91.7	541	7	ADC07369	Human	sec
590	22	91.7	541	6	ABR72478	Human	sec	663	22	91.7	541	7	ADC11359	Human	sec
591	22	91.7	541	6	ABR74308	Human	sec	664	22	91.7	541	7	ADC14481	Novel	hum
592	22	91.7	541	6	ABO18576	Human	sec	665	22	91.7	541	7	ADD08013	Novel	hum
593	22	91.7	541	6	ABR80282	Human	sec	666	22	91.7	541	7	ADD81838	Human	PRO
594	22	91.7	541	6	ABM01503	Human	sec	667	22	91.7	541	7	ADD07480	Novel	hum
595	22	91.7	541	6	ABM02113	Human	sec	668	22	91.7	541	7	ADC82371	Human	PRO
596	22	91.7	541	6	ABR87275	Human	sec	669	22	91.7	541	7	ADD05625	Human	sec
597	22	91.7	541	6	ABM12842	Human	sec	670	22	91.7	541	7	ADD08551	Novel	hum
598	22	91.7	541	6	ABM30596	Human	sec	671	22	91.7	541	7	ADD06800	Novel	hum
599	22	91.7	541	6	ABM24496	Human	sec	672	22	91.7	541	7	ADD83047	Human	PRO
600	22	91.7	541	6	ABO29410	Human	sec	673	22	91.7	541	7	ADD55154	Human	PRO
601	22	91.7	541	6	ABO31240	Human	sec	674	22	91.7	541	7	ADD56112	Human	PRO
602	22	91.7	541	6	ABM14367	Human	sec	675	22	91.7	541	7	ADD54550	Human	PRO
603	22	91.7	541	6	ABM09792	Human	sec	676	22	91.7	541	7	ADE26704	Novel	hum
604	22	91.7	541	6	ABO38917	Human	sec	677	22	91.7	541	7	ADE26171	Novel	hum
605	22	91.7	541	6	ABM34682	Human	sec	678	22	91.7	541	7	ADF67108	Human	PRO
606	22	91.7	541	6	ABO51158	Human	sec	679	22	91.7	541	7	ADG02620	Novel	hum
607	22	91.7	541	6	ABO03984	Human	sec	680	22	91.7	541	7	ADG01327	Novel	hum
608	22	91.7	541	6	ABO10454	Human	PRO	681	22	91.7	541	7	ADF95502	Novel	hum
609	22	91.7	541	6	ABO53150	Human	sec	682	22	91.7	541	7	ADG12317	Novel	hum
	22	91.7	541	6	ABR77697	Human	sec								
	22	91.7	541	6	ABR78907	Human	sec								

683	22	91.7	541	7	ADH08977	Adh08977 Human PRO	756	22	91.7	762	5	AAE22116	Aae22116 Desulfuro
684	22	91.7	541	7	ADI35362	Adi35362 Human PRO	757	22	91.7	765	6	ABU31950	Abu31950 Protein e
685	22	91.7	541	7	ADH99854	Adh99854 Novel hum	758	22	91.7	775	6	ADA55050	Ada55050 Human pro
686	22	91.7	541	7	ADL32758	Adl32758 Novel hum	759	22	91.7	783	5	ABP73575	Abp73575 Candida a
687	22	91.7	541	7	ADM30292	Adm30292 Novel hum	760	22	91.7	784	8	ADN46684	Adn46684 Thermococ
688	22	91.7	541	7	ADe74289	AdE74289 Human sec	761	22	91.7	789	5	ABP41761	Abp41761 Human ova
689	22	91.7	541	8	ADe74901	AdE74901 Human sec	762	22	91.7	805	6	ADA27366	Ada27366 HPV-16 L2
690	22	91.7	541	8	ADe74901	AdE74901 Human PRO	763	22	91.7	805	6	ADA27366	Ada27366 HPV-16 L2
691	22	91.7	541	8	ADG11557	Adg11557 Human PRO	764	22	91.7	805	7	ADA14293	Ada14293 HPV-16 L2
692	22	91.7	541	8	ADG04385	Adg04385 Novel hum	765	22	91.7	805	7	ADAE38618	Aae38618 HPV-16 L2
693	22	91.7	541	8	ADG00545	Adg00545 Novel hum	766	22	91.7	815	7	ABO62060	AbO62060 Klebsiell
694	22	91.7	541	8	ADG00545	Adg00545 Novel hum	767	22	91.7	819	6	ABM68990	Abm68990 Photorhab
695	22	91.7	541	8	ADG82801	Adg82801 Human PRO	768	22	91.7	821	8	ADM45830	Adm45830 TNF-alpha
696	22	91.7	541	8	ADH26082	Adh26082 Novel hum	769	22	91.7	825	6	ABJ25874	Abj25874 Aspergill
697	22	91.7	541	8	ADH19427	Adh19427 Human sec	770	22	91.7	825	6	ABU50425	Abu50425 Protein e
698	22	91.7	541	8	ADH33051	Adh33051 Human PRO	771	22	91.7	827	4	AAV72134	Aav72134 E. coli g
699	22	91.7	541	8	ADH20920	Adh20920 Human sec	772	22	91.7	827	4	AAV72133	Aav72133 Escherich
700	22	91.7	541	8	ADH19960	Adh19960 Human sec	773	22	91.7	858	6	ABR53197	AbR53197 Protein s
701	22	91.7	541	8	ADJ54790	Adj54790 Human PRO	774	22	91.7	858	7	ABJ26474	Abj26474 Aspergill
702	22	91.7	541	8	ADJ64561	Adj64561 Human PRO	775	22	91.7	872	6	ABJ26474	Abj26474 Aspergill
703	22	91.7	541	8	ADM31457	Adm31457 Novel hum	776	22	91.7	880	4	ABB71719	Abb71719 Drosophi
704	22	91.7	541	8	ADM36504	Adm36504 Novel hum	777	22	91.7	881	7	ABO65486	AbO65486 Klebsiell
705	22	91.7	541	8	ADN37917	Adn37917 Novel hum	778	22	91.7	885	4	ABN61502	Abn61502 Drosophi
706	22	91.7	541	8	ADN34095	Adn34095 Staphyloc	779	22	91.7	885	6	ABM04826	Abm04826 Human KIA
707	22	91.7	541	8	ABU30574	Abu30574 Protein e	780	22	91.7	900	2	AAO65454	AaO65454 Viral enh
708	22	91.7	552	2	ARI3879	Aar13879 Indoleace	781	22	91.7	901	2	AAO65454	AaO65454 Viral enh
709	22	91.7	553	4	AAU36605	Aau36605 Staphyloc	782	22	91.7	901	2	AAO65454	AaO65454 Viral enh
710	22	91.7	553	6	ABU16137	Abu16137 Protein e	783	22	91.7	901	2	AAO65454	AaO65454 Viral enh
711	22	91.7	553	6	ABM72931	Abm72931 Staphyloc	784	22	91.7	915	4	ABB12073	Abb12073 Human sec
712	22	91.7	555	6	ADA54711	Ada54711 Human pro	785	22	91.7	915	4	ABB12073	Abb12073 Human sec
713	22	91.7	555	6	ABP39547	Abp39547 Staphyloc	786	22	91.7	915	4	ABB12073	Abb12073 Human sec
714	22	91.7	557	5	ABU30574	Abu30574 Protein e	787	22	91.7	915	4	ABB12073	Abb12073 Human sec
715	22	91.7	557	6	ABU30574	Abu30574 Protein e	788	22	91.7	915	4	ABB12073	Abb12073 Human sec
716	22	91.7	564	3	AAO65454	AaO65454 Viral enh	789	22	91.7	915	4	ABB12073	Abb12073 Human sec
717	22	91.7	564	3	AAO65454	AaO65454 Viral enh	790	22	91.7	915	4	ABB12073	Abb12073 Human sec
718	22	91.7	565	4	ABU76817	Abu76817 Coryneb	791	22	91.7	915	4	ABB12073	Abb12073 Human sec
719	22	91.7	576	5	ABY25712	AbY25712 Rat prote	792	22	91.7	1015	8	ADU50405	AdU50405 Oil-ssoc
720	22	91.7	578	3	ABN28675	Abn28675 Human car	793	22	91.7	1072	8	ADU50405	AdU50405 Oil-ssoc
721	22	91.7	578	4	AAW40317	AaW40317 Human pol	794	22	91.7	1075	5	ABP74097	Abp74097 Human TRI
722	22	91.7	578	4	ABN55598	Abn55598 Human pro	795	22	91.7	1081	6	ADA55460	Ada55460 Human pro
723	22	91.7	578	4	ADA55598	Ada55598 Human pro	796	22	91.7	1081	6	ADA55460	Ada55460 Human pro
724	22	91.7	578	6	ADA56824	Ada56824 Human sec	797	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
725	22	91.7	578	6	ADA56824	Ada56824 Human sec	798	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
726	22	91.7	578	7	ADC74067	Adc74067 Human sec	799	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
727	22	91.7	578	7	ADD37854	AdD37854 Human pol	800	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
728	22	91.7	579	4	AAW43527	AaW43527 Human pol	801	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
729	22	91.7	579	4	AAW43527	AaW43527 Human pol	802	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
730	22	91.7	579	8	ADM24548	Adm24548 Human PRO	803	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
731	22	91.7	582	6	ABM72067	Abm72067 Staphyloc	804	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
732	22	91.7	583	7	ADC97184	Adc97184 E. faeciu	805	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
733	22	91.7	589	7	ADJ65849	Adj65849 C. glutam	806	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
734	22	91.7	591	6	ADA54983	Ada54983 Human pro	807	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
735	22	91.7	616	4	AAW42103	AaW42103 Human pol	808	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
736	22	91.7	616	4	AAW42103	AaW42103 Human pol	809	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
737	22	91.7	627	2	AAW85597	AaW85597 Human GAB	810	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
738	22	91.7	627	6	ABU19813	AbU19813 Androgen-	811	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
739	22	91.7	627	6	ABU19117	AbU19117 Protein e	812	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
740	22	91.7	632	6	ABU19117	AbU19117 Protein e	813	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
741	22	91.7	632	6	ABU32823	AbU32823 Protein e	814	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
742	22	91.7	632	6	ABU32823	AbU32823 Protein e	815	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
743	22	91.7	635	6	ABU22011	Aay68835 The poly(	816	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
744	22	91.7	637	3	AAV68835	Abj49360 Oil-ssoc	817	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
745	22	91.7	637	3	AAV68835	Abj49360 Oil-ssoc	818	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
746	22	91.7	638	8	ADJ49360	Adj49360 Protein e	819	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
747	22	91.7	639	6	ABU19847	Abu19847 Protein e	820	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
748	22	91.7	704	6	ADA34250	Ada34250 Acinetoba	821	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
749	22	91.7	717	4	ABG17809	Abg17809 Novel hum	822	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
750	22	91.7	717	4	ABG18220	Abg18220 Novel hum	823	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
751	22	91.7	720	7	ADJ70357	Adj70357 Human hea	824	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
752	22	91.7	726	6	ABU24278	Abu24278 Protein e	825	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
753	22	91.7	740	7	ADE31208	Ade31208 Human dia	826	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
754	22	91.7	761	4	ABA64719	AbA64719 T. brucei	827	22	91.7	1090	4	ABG17694	Abg17694 Novel hum
755	22	91.7	762	2	AAV00939	Aay00939 Desulfuro	828	22	91.7	1090	4	ABG17694	Abg17694 Novel hum

829	22	91.7	1873	8	ADJ32536	Adj32536	Human int	902	21	87.5	79	4	ABG22403	Abg22403	Novel hum
830	22	91.7	1873	8	ADJ32525	Adj32525	Human int	903	21	87.5	81	5	ABP01866	Abp01866	Human ORF
831	22	91.7	1875	4	AAB68089	Aam68089	Amino aci	904	21	87.5	84	4	AAM21351	Aam21351	Peptide #
832	22	91.7	1978	2	AYAY07032	AYay07032	Breast ca	905	21	87.5	84	4	ABB43688	Abb43688	Peptide #
833	22	91.7	1978	6	ABR64250	Abr64250	Angiogene	906	21	87.5	84	4	AAM37583	Aam37583	Peptide #
834	22	91.7	2208	8	ADN72815	Adn72815	Thale cre	907	21	87.5	84	4	ABM26636	Abm26636	Protein #
835	22	91.7	3892	6	ADA34216	Ada34216	Acinetoba	908	21	87.5	84	4	AAM64647	Aam64647	Human bra
836	22	91.7	5701	4	ABB36684	Abb36684	Peptide #	909	21	87.5	84	4	ABG59061	Abg59061	Human liv
837	22	91.7	5701	4	ABB22021	Abb22021	Protein #	910	21	87.5	86	4	AAB46501	Aab46501	B. subtil
838	22	91.7	5701	4	AAM69843	Aam69843	Human bon	911	21	87.5	86	4	AAB46483	Aab46483	B. lichen
839	22	91.7	5701	4	AAM57448	Aam57448	Human bra	912	21	87.5	86	4	AAB46498	Aab46498	B. subtil
840	22	91.7	5701	4	ARG51536	Arg51536	Human liv	913	21	87.5	86	4	ABG07436	Abg07436	Novel hum
841	22	91.7	5701	4	RAM05328	Ram05328	Peptide #	914	21	87.5	86	5	ABG69477	Abg69477	Human bai
842	22	91.7	5701	8	ABO58804	Abo58804	Human gen	915	21	87.5	87	6	AAU58059	Aau58059	Propionib
843	22	91.7	5701	4	AAU05396	Aau05396	Human tit	916	21	87.5	87	6	ABM54578	Abm54578	Propionib
844	22	91.7	26926	8	ADQ17316	Adq17316	Human sof	917	21	87.5	88	8	ABO60345	AbO60345	Human gen
845	22	91.7	26926	6	AAAB31886	Aab31886	Peptide f	918	21	87.5	90	4	AAM21427	Aam21427	Peptide #
846	21	87.5	6	AAAB31885	Aab31885	Peptide f	919	21	87.5	90	4	ABB43765	Abb43765	Peptide #	
847	21	87.5	9	ADM12723	Adm12723	MHC class	920	21	87.5	90	4	RAM37668	Ram37668	Protein #	
848	21	87.5	9	ADO38960	Ado38960	Anthrax p	921	21	87.5	90	4	ABB26708	Abb26708	Protein #	
849	21	87.5	10	AAG96723	Aag96723	Human com	922	21	87.5	90	4	AAM77498	Aam77498	Human bon	
850	21	87.5	10	AAG85894	Aag85894	Saccharom	923	21	87.5	90	4	AAM64727	Aam64727	Human bra	
851	21	87.5	10	AAG85892	Aag85892	Saccharom	924	21	87.5	90	4	ABG59136	Abg59136	Human liv	
852	21	87.5	10	ADK05725	Adk05725	Hepatitis	925	21	87.5	90	5	ADK36999	Adk36999	Novel hum	
853	21	87.5	10	ADK05726	Adk05726	Hepatitis	926	21	87.5	91	5	ABP06817	Abp06817	Human ORF	
854	21	87.5	11	AAU82656	Aau82656	BP protei	927	21	87.5	91	5	ABP26163	Abp26163	Streptoco	
855	21	87.5	11	AAU82657	Aau82657	BP protei	928	21	87.5	92	5	ABU67208	Abu67208	G-protein	
856	21	87.5	12	AAR67199	Aar67199	Bovine gl	929	21	87.5	92	6	ABP53989	Abp53989	Human bet	
857	21	87.5	12	AAR67199	Aar67199	Bovine gl	929	21	87.5	92	6	ADO28729	Ado28729	Human bet	
858	21	87.5	12	AAR66663	Aar66663	GGF II tr	930	21	87.5	92	8	ADO05233	Ado05233	Beta-2-ad	
859	21	87.5	16	AAU07813	Aau07813	Peptide i	931	21	87.5	92	8	ADO05233	Ado05233	Beta-2-ad	
860	21	87.5	18	AAAB20421	Abab20421	Anti-FIX/	932	21	87.5	94	5	AAE13241	Aae13241	Human lin	
861	21	87.5	18	ABP82141	Abp82141	G protein	933	21	87.5	95	3	AAE13241	Aae13241	Human lin	
862	21	87.5	21	AAW67046	Aaw67046	Laminin p	934	21	87.5	97	8	ADN48125	Adn48125	Thermoco	
863	21	87.5	21	AAW67045	Aaw67045	Laminin p	935	21	87.5	98	4	ABB28808	Abb28808	Peptide #	
864	21	87.5	21	ABR42890	Abr42890	Bovine hi	936	21	87.5	98	4	ABBI19433	Abbi19433	Protein #	
865	21	87.5	25	ADM12746	Adm12746	Ii key/AR	937	21	87.5	98	4	ABM54758	Abm54758	Human bra	
866	21	87.5	25	ADO38983	Ado38983	Anthrax p	938	21	87.5	98	4	ABG48812	Abg48812	Human liv	
867	21	87.5	26	AAG99693	Aag99693	ERA bindi	939	21	87.5	99	3	AGG34449	Agg34449	Arabidops	
868	21	87.5	28	ADP83926	Adp83926	Human BAC	940	21	87.5	100	3	AAG57503	Agg57503	Arabidops	
869	21	87.5	37	AAAM33996	Aam33996	Peptide #	941	21	87.5	100	6	ABU70559	Abu70559	Human adi	
870	21	87.5	37	AAAM73810	Aam73810	Human bon	942	21	87.5	100	7	ADB64208	Adb64208	Human pro	
871	21	87.5	37	AAW61103	Aaw61103	Human bra	943	21	87.5	101	8	ABB23046	Abb23046	Protein #	
872	21	87.5	37	ABG55558	Abg55558	Human liv	944	21	87.5	101	8	ABO55363	AbO55363	Human gen	
873	21	87.5	37	ABG43697	Abg43697	Human pep	945	21	87.5	102	2	AAW78901	Aaw78901	Human UNC	
874	21	87.5	38	ADK01884	Adk01884	Hepatitis	946	21	87.5	102	4	AAM30999	Aam30999	Peptide #	
875	21	87.5	47	AAAG75334	Aag75334	Human col	947	21	87.5	102	4	ABB22872	Abb22872	Protein #	
876	21	87.5	47	AAAM2985	Aam2985	Human imm	948	21	87.5	102	8	ADK16101	Adk16101	Nanoarcha	
877	21	87.5	49	AAAY55690	Aay55690	S. pombe	949	21	87.5	103	2	AAAY34067	Aay34067	Histone H	
878	21	87.5	49	AAAY5679	Aay5679	E. coli y	950	21	87.5	103	2	AAAY57365	Aay57365	Human his	
879	21	87.5	49	AAAY55680	Aay55680	H. influe	951	21	87.5	103	7	ADC88510	Adc88510	Ribosomal	
880	21	87.5	49	ABB17187	Abb17187	Human ner	952	21	87.5	104	2	AAAR54312	Aar54312	Anti-HIV	
881	21	87.5	50	ABP06211	Abp06211	Human ORF	953	21	87.5	104	3	AAAY55837	Aay55837	Mouse SPR	
882	21	87.5	51	ABR42893	Abr42893	Bovine hi	954	21	87.5	105	3	AAG01392	Aag01392	Human sec	
883	21	87.5	53	AAAR43265	Aar43265	RGAL C-te	955	21	87.5	107	7	ADC96202	Adc96202	E. faeciu	
884	21	87.5	54	AAAG99655	Aag99655	ERA bindi	956	21	87.5	108	4	ABG07141	Abg07141	Novel hum	
885	21	87.5	58	AAAG4956	Aag4956	Zea mays	957	21	87.5	109	7	ADH87878	Adh87878	Enterococ	
886	21	87.5	58	AAAG0195	Aag0195	C glutami	958	21	87.5	110	5	ADII17270	Adi17270	Polypepti	
887	21	87.5	59	ABP02868	Abp02868	Human ORF	959	21	87.5	110	5	ADII17270	Adi17270	Polypepti	
888	21	87.5	62	AAAY48284	Aay48284	Human pro	960	21	87.5	111	4	ABG07149	Abg07149	Novel hum	
889	21	87.5	62	AAU23633	Aau23633	Novel hum	961	21	87.5	112	3	ABAB1337	Abab1337	Human ORF	
890	21	87.5	62	ABR42894	Abr42894	Bovine hi	962	21	87.5	112	6	ABO27167	AbO27167	Humanisee	
891	21	87.5	62	ABO58104	AbO58104	Human gen	963	21	87.5	112	6	ABO27167	AbO27167	Humanisee	
892	21	87.5	64	ABP09204	Abp09204	Human ORF	964	21	87.5	114	3	AAG44954	Aag44954	Zea mays	
893	21	87.5	64	ABR55391	Abr55391	Amino aci	965	21	87.5	114	4	ABB67449	Abb67449	Drosophil	
894	21	87.5	64	ADM26846	Adm26846	Hyperther	966	21	87.5	114	5	ABP08643	Abp08643	Human ORF	
895	21	87.5	65	ABU23965	Abu23965	Protein e	967	21	87.5	115	3	RAY28420	Ray28420	Corn bran	
896	21	87.5	66	ABR55388	Abr55388	Amino aci	968	21	87.5	115	4	AAAM96473	Aam96473	Human rep	
897	21	87.5	68	ADA48192	Ada48192	Rice prot	969	21	87.5	115	4	ABG07133	Abg07133	Novel hum	
898	21	87.5	69	ABJ04855	Abj04855	Human pro	970	21	87.5	115	5	ABBA49101	Abb49101	Listeria	
899	21	87.5	70	AAW92635	Aaw92635	D. melano	971	21	87.5	116	2	AAAY34066	Aay34066	Histone H	
900	21	87.5	75	AAAY11284	Aay11284	S. pneumo	972	21	87.5	116	2	AAAY57364	Aay57364	Human his	
901	21	87.5	75	AAO06283	Aao06283	Human pol	973	21	87.5	116	5	ABBA49265	Abb49265	Listeria	
902	21	87.5	76	ABP09594	Abp09594	Human ORF	974	21	87.5	118	4	AAAB48256	Aab48256	Corn magn	

975 21 87.5 118 4 AAU41271 Aau41271 Propionib  
976 21 87.5 118 6 ABM37790 Abm37790 Propionib  
977 21 87.5 118 6 ABO09050 Abo09050 Alloiococ  
978 21 87.5 120 4 ABG13644 Abg13644 Novel hum  
979 21 87.5 120 4 ABG13648 Abg13648 Novel hum  
980 21 87.5 120 7 ADI53296 Adi53296 Human his  
981 21 87.5 121 4 AAU37243 Aau37243 Staphyloc  
982 21 87.5 121 4 AAU34049 Aau34049 Staphyloc  
983 21 87.5 121 4 AAU36636 Aau36636 Staphyloc  
984 21 87.5 121 5 AAO20062 Aao20062 Staphyloc  
985 21 87.5 121 5 AAO20083 Aao20083 Staphyloc  
986 21 87.5 121 6 ABU42606 Abu42606 Protein e  
987 21 87.5 121 6 ABU16254 Abu16254 Protein e  
988 21 87.5 121 6 ABM73409 Abm73409 Staphyloc  
989 21 87.5 122 3 AAG57502 Aag57502 Arabiodops  
990 21 87.5 122 4 ABG11471 Abg11471 Novel hum  
991 21 87.5 123 5 ABP38414 Abp38414 Staphyloc  
992 21 87.5 124 8 ADN99599 Adn99599 Novel hum  
993 21 87.5 124 8 ADP09946 Adp09946 Anti-mous  
994 21 87.5 124 8 ADO25696 Ado25696 Anti-TNF-  
995 21 87.5 125 4 ABG26639 Abg26639 Novel hum  
996 21 87.5 126 8 ADK48660 Adk48660 Streptoco  
997 21 87.5 126 8 ADP29338 Adp29338 Human sec  
998 21 87.5 127 2 AAR74764 Aar74764 Sermun am  
999 21 87.5 128 5 ABB54819 Abb54819 Lactococc  
1000 21 87.5 128 7 ADH85685 Adh85685 Enterococ

ALIGNMENTS

RESULT 1  
AAR15552  
ID AAR15552 standard; protein; 20 AA.  
AC AAR15552;  
XX  
XX  
DT 02-MAR-1992 (first entry)  
XX  
XX Immunopeptide #3 derived from HPV16 E2 peptide.  
DE cervical cancer; cervical intraepithelial neoplasia; CIN; wart;  
KW squamous cell carcinoma; ELISA; HPV 16.  
XX  
XX Synthetic.  
OS  
XX WO9118294-A.  
PN  
XX  
XX  
PD 28-NOV-1991.  
XX  
XX  
PF 11-MAY-1990; 90SE-00001705.  
XX  
XX 11-MAY-1990; 90SE-00001705.  
PR  
XX  
XX (MEDS-) MEDSCAND AB.  
PA  
XX Dillner J, Dillner L, Cheng HM;  
PI  
XX WPI; 1991-369390/50.  
DR  
XX  
XX Diagnosis of human papilloma virus infection and PV-carrying tumours -  
PT using synthetic peptide(s) to detect virus specific antigen-antibody  
PT complexes by immunoassay.  
XX  
XX  
PS Disclosure; Page 38; 72pp; English.  
XX  
XX This is one of a large number of peptides which have been synthesised on  
CC the basis of the amino acid sequences for the E2, E4, E7, L1 or L2  
CC proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The selection of  
CC peptide sequences was based on the assumption that an immunoreactive  
CC region might be situated in the same relative region of a protein from  
CC different HPV types. The peptides were used in diagnostic immunoassays to  
CC detect HPV-infection. See AAR15523-RU5601

XX  
SQ Sequence 20 AA;  
XX  
XX Query Match 91.7%; Score 22; DB 2; Length 20;  
Best Local Similarity 80.0%; Pred. No. 1e+02; 1; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 1;  
QY 1 GYXVE 5  
Db 9 GYIVE 13  
XX  
XX RESULT 2  
ADF70007  
ID ADF70007 standard; protein; 45 AA.  
XX  
XX ADF70007;  
AC  
XX  
DT 12-FEB-2004 (first entry)  
XX  
XX AcmaA-type homologue amino acid sequence.  
DE  
XX delivery; targeting system; AcmaA-type anchor protein; solid tumour;  
KW health; medical; agricultural; cosmetic; controlled release.  
XX  
XX Lactococcus lactis.  
OS  
XX WO2003084508-A1.  
PN  
XX  
XX 16-OCT-2003.  
PD  
XX  
XX 04-APR-2003; 2003WO-NL000256.  
PF  
XX  
XX 04-APR-2002; 2002EP-00076316.  
PR  
XX 04-APR-2002; 2002US-0369927P.  
PR  
XX 05-APR-2002; 2002US-0370485P.  
PR  
XX 20-DEC-2002; 2002EP-00080481.  
PR  
XX  
XX (NANO-) APPLIED NANOSYSTEMS BV.  
PA  
XX  
XX Friesen RHE, Leenhouts CJ, Hektor HJ, Van Esch JH, Heeres A;  
PI Robillard GT;  
PI  
XX  
XX WPI; 2003-877005/81.  
DR  
XX  
XX Vehicle for targeted delivery of therapeutic or diagnostic agents,  
PT includes protein anchor and system for inducing availability of the  
PT agent.  
XX  
XX  
XX Example 3; Page 191; 303pp; English.  
PS  
XX  
XX The present invention describes a vehicle (A) for delivering a substance  
CC (I) to a predetermined site, which comprises (I); a system for inducing  
CC availability of at least one compartment of (A) towards the exterior;  
CC and, as targeting system for directing (A) to the site, an AcmaA-type  
CC anchor protein (II). (A) are used for delivery of diagnostic and  
CC therapeutic agents to predetermined sites in the body, particularly  
CC joints or solid tumours but can be used more generally for health,  
CC medical, agricultural and cosmetic applications. (A) significantly  
CC increases the half-life of peptides in the circulation and, by providing  
CC controlled release, ensures relatively high bioavailability, allowing  
CC therapeutic use of agents that would otherwise be too toxic for systemic  
CC administration. The native AcmaA peptide targets gram-positive bacteria  
CC but its homologues can be engineered to have different selectivity. The  
CC present sequence is used in the exemplification of the present invention.  
XX  
XX Sequence 45 AA;  
SQ  
XX  
XX Query Match 91.7%; Score 22; DB 7; Length 45;  
Best Local Similarity 80.0%; Pred. No. 2.4e+02; 1; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 1;  
QY 1 GYXVE 5

```

Db      || ||
      18 GYSVE 22

RESULT 3
AAW75574
ID  AAW75574 standard; protein; 46 AA.
XX
AC  AAW75574;
XX
DT  23-OCT-1998 (first entry)
XX
DE  M. tuberculosis 16 kD extracellular product N-terminal sequence 2.
XX
KW  Mycobacterium tuberculosis; vaccination; extracellular product;
KW  immunodominant epitope; interleukin-12; MF59; immune response;
KW  opsonising humoral response; intracellular pathogen.
XX
OS  Mycobacterium tuberculosis.
XX
FN  WO9831388-A1.
XX
PD  23-JUL-1998.
XX
PF  15-JAN-1998; 98WO-US000942.
XX
PR  21-JAN-1997; 97US-00786533.
XX
PA  (REGC ) UNIV CALIFORNIA.
XX
PI  Horwitz MA, Harth G, Lee B;
XX
DR  WPI; 1998-413815/35.
XX
XX  Vaccines against Mycobacterium containing major extracellular proteins -
PT  used to, e.g. induce protective and therapeutic immune responses, and for
PT  detecting an immune response.
XX
PS  Example 2; Page 35; 236pp; English.
XX
CC  Sequences shown in AAW75570 to AAW75586 represent N-terminal amino acid
CC  sequences of 14 exemplary major abundant extracellular products of
CC  Mycobacterium tuberculosis. The invention provides an agent for
CC  vaccinating mammals against Mycobacterium. The agent comprises at least
CC  one of the major abundant extracellular 110, 80, 71, 58, 45, 32A, 32B,
CC  30, 24, 23.5, 23, 16, 14 or 12 kDa proteins of M. tuberculosis, or at
CC  least 1 of their immunodominant epitopes and interleukin-12 (IL-12) or
CC  MF59 as adjuvants. The agent containing the nucleic acid encoding the
CC  extracellular products are used to raise a protective or therapeutic
CC  immune response against Mycobacterium, specifically M. tuberculosis. The
CC  immunodominant epitopes can also be used (typically in a cutaneous
CC  hypersensitivity test) to detect an immune response to vaccination.
CC  Preparation of the agent does not require selection of the most
CC  immunogenic products, so large scale production and purification are
CC  easy, resulting in a consistent, standardised formulation, having lower
CC  toxicity than killed or attenuated vaccines. The agents provide a rapid
CC  and effective response (including a strong cell-mediated component) and
CC  are safe even in immunocompromised subjects. They prevent development of
CC  an opsonising humoral response that might spread intracellular pathogens
XX
SQ  Sequence 46 AA;

Query Match      91.7%; Score 22; DB 2; Length 46;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 GYXVE 5
Db      || ||
      39 GYTVE 43

RESULT 4
AAW75573
ID  AAW75573 standard; protein; 46 AA.
XX
AC  AAW75573;
XX
DT  23-OCT-1998 (first entry)
XX
DE  M. tuberculosis 16 kD extracellular product N-terminal sequence 1.
XX
KW  Mycobacterium tuberculosis; vaccination; extracellular product;
KW  immunodominant epitope; interleukin-12; MF59; immune response;
KW  opsonising humoral response; intracellular pathogen.
XX
OS  Mycobacterium tuberculosis.
XX
FN  WO9831388-A1.
XX
PD  23-JUL-1998.
XX
PF  15-JAN-1998; 98WO-US000942.
XX
PR  21-JAN-1997; 97US-00786533.
XX
PA  (REGC ) UNIV CALIFORNIA.
XX
PI  Horwitz MA, Harth G, Lee B;
XX
DR  WPI; 1998-413815/35.
XX
XX  Vaccines against Mycobacterium containing major extracellular proteins -
PT  used to, e.g. induce protective and therapeutic immune responses, and for
PT  detecting an immune response.
XX
PS  Example 2; Page 35; 236pp; English.
XX
CC  Sequences shown in AAW75570 to AAW75586 represent N-terminal amino acid
CC  sequences of 14 exemplary major abundant extracellular products of
CC  Mycobacterium tuberculosis. The invention provides an agent for
CC  vaccinating mammals against Mycobacterium. The agent comprises at least
CC  one of the major abundant extracellular 110, 80, 71, 58, 45, 32A, 32B,
CC  30, 24, 23.5, 23, 16, 14 or 12 kDa proteins of M. tuberculosis, or at
CC  least 1 of their immunodominant epitopes and interleukin-12 (IL-12) or
CC  MF59 as adjuvants. The agent containing the nucleic acid encoding the
CC  extracellular products are used to raise a protective or therapeutic
CC  immune response against Mycobacterium, specifically M. tuberculosis. The
CC  immunodominant epitopes can also be used (typically in a cutaneous
CC  hypersensitivity test) to detect an immune response to vaccination.
CC  Preparation of the agent does not require selection of the most
CC  immunogenic products, so large scale production and purification are
CC  easy, resulting in a consistent, standardised formulation, having lower
CC  toxicity than killed or attenuated vaccines. The agents provide a rapid
CC  and effective response (including a strong cell-mediated component) and
CC  are safe even in immunocompromised subjects. They prevent development of
CC  an opsonising humoral response that might spread intracellular pathogens
XX
SQ  Sequence 46 AA;

Query Match      91.7%; Score 22; DB 2; Length 46;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 GYXVE 5
Db      || ||
      39 GYTVE 43

RESULT 5
AAW75573
ID  AAW75573 standard; peptide; 47 AA.
XX
AC  AAW75573;
XX
DT  30-SEP-1996 (first entry)
XX

```

```

ID  AAW75573 standard; protein; 46 AA.
XX
AC  AAW75573;
XX
DT  23-OCT-1998 (first entry)
XX
DE  M. tuberculosis 16 kD extracellular product N-terminal sequence 1.
XX
KW  Mycobacterium tuberculosis; vaccination; extracellular product;
KW  immunodominant epitope; interleukin-12; MF59; immune response;
KW  opsonising humoral response; intracellular pathogen.
XX
OS  Mycobacterium tuberculosis.
XX
FN  WO9831388-A1.
XX
PD  23-JUL-1998.
XX
PF  15-JAN-1998; 98WO-US000942.
XX
PR  21-JAN-1997; 97US-00786533.
XX
PA  (REGC ) UNIV CALIFORNIA.
XX
PI  Horwitz MA, Harth G, Lee B;
XX
DR  WPI; 1998-413815/35.
XX
XX  Vaccines against Mycobacterium containing major extracellular proteins -
PT  used to, e.g. induce protective and therapeutic immune responses, and for
PT  detecting an immune response.
XX
PS  Example 2; Page 35; 236pp; English.
XX
CC  Sequences shown in AAW75570 to AAW75586 represent N-terminal amino acid
CC  sequences of 14 exemplary major abundant extracellular products of
CC  Mycobacterium tuberculosis. The invention provides an agent for
CC  vaccinating mammals against Mycobacterium. The agent comprises at least
CC  one of the major abundant extracellular 110, 80, 71, 58, 45, 32A, 32B,
CC  30, 24, 23.5, 23, 16, 14 or 12 kDa proteins of M. tuberculosis, or at
CC  least 1 of their immunodominant epitopes and interleukin-12 (IL-12) or
CC  MF59 as adjuvants. The agent containing the nucleic acid encoding the
CC  extracellular products are used to raise a protective or therapeutic
CC  immune response against Mycobacterium, specifically M. tuberculosis. The
CC  immunodominant epitopes can also be used (typically in a cutaneous
CC  hypersensitivity test) to detect an immune response to vaccination.
CC  Preparation of the agent does not require selection of the most
CC  immunogenic products, so large scale production and purification are
CC  easy, resulting in a consistent, standardised formulation, having lower
CC  toxicity than killed or attenuated vaccines. The agents provide a rapid
CC  and effective response (including a strong cell-mediated component) and
CC  are safe even in immunocompromised subjects. They prevent development of
CC  an opsonising humoral response that might spread intracellular pathogens
XX
SQ  Sequence 46 AA;

Query Match      91.7%; Score 22; DB 2; Length 46;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 GYXVE 5
Db      || ||
      39 GYTVE 43

RESULT 5
AAW75573
ID  AAW75573 standard; peptide; 47 AA.
XX
AC  AAW75573;
XX
DT  30-SEP-1996 (first entry)
XX

```

DE Mycobacterium 16 kD protein N-terminal peptide.  
 XX N-terminal peptide; extracellular product; Mycobacterium; 16 kD protein;  
 KW immunoreactive; vaccine; immune response; pathogen; mammal; therapy;  
 KW opsonising humoral response; intracellular bacteria.  
 XX Mycobacterium tuberculosis.  
 OS  
 XX WO9605223-A1.  
 XX  
 XX 22-FEB-1996.  
 XX  
 XX 24-FEB-1995; 95WO-US002373.  
 XX  
 XX 12-AUG-1994; 94US-00289667.  
 XX  
 XX (REGC ) UNIV CALIFORNIA.  
 XX  
 XX Horwitz MA;  
 XX  
 XX WPI; 1996-139640/14.  
 XX  
 XX Vaccines comprising abundant extracellular prods. - useful to vaccinate  
 PT against intracellular bacteria, partic. Mycobacterium tuberculosis.  
 PT  
 XX Claim 60; Page 109; 139pp; English.  
 PS  
 XX AA92876-R92889 represent N-terminal fragments of the abundant  
 CC extracellular products of Mycobacterium. This sequence represents the N-  
 CC terminus of the 16 kD protein of M. tuberculosis. The 16 kD protein, or  
 CC an immunoreactive homologue or a fragment of it, with this sequence as  
 CC the N-terminus can be used in a vaccinating agent of the invention. The  
 CC vaccinating agent promotes an immune response against a Mycobacterium  
 CC pathogen, and are therefore useful for immunising a mammal against one of  
 CC these pathogens. As the vaccines do not contain whole bacteria or  
 CC components, they are less likely to provoke toxic responses compared to  
 CC attenuated or killed bacterial vaccines. Also, the vaccines are not life-  
 CC threatening to immunocompromised individuals. The use of extracellular  
 CC products also prevents the development of an opsonising humoral response  
 CC which can increase the pathogenesis of intracellular bacteria. It also  
 CC reduces the potential for generating a response which precludes the use  
 CC of widely used screening and control techniques based on host recognition  
 CC of immunogenic agents  
 XX  
 XX Sequence 47 AA;  
 SQ  
 Query Match 91.7%; Score 22; DB 2; Length 47;  
 Best Local Similarity 80.0%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 40 GYTVE 44  
 RESULT 6  
 ABG74428  
 ID ABG74428 standard; peptide; 47 AA.  
 AC  
 XX ABG74428;  
 XX  
 XX 14-APR-2003 (first entry)  
 DT  
 XX M. tuberculosis extended extracellular protein N-terminal peptide #13.  
 DE  
 XX Tuberculosis; extracellular protein; antibacterial; tuberculostatic;  
 KW antileprotic; virucide; fungicide; protozoacide; vaccine; virus;  
 KW immune response; Mycobacterium; infectious pathogen; protozoa; fungi;  
 KW bacteria.  
 XX  
 XX Mycobacterium tuberculosis.  
 OS  
 XX US2002150592-A1.  
 XX , PN

XX 17-OCT-2002.  
 PD  
 XX 14-SEP-2001; 2001US-00953457.  
 PF  
 XX 18-SEP-1998; 98US-00156358.  
 PR  
 XX (REGC ) UNIV CALIFORNIA.  
 XX  
 XX Horwitz MA;  
 XX  
 XX WPI; 2003-198266/19.  
 XX  
 XX New vaccinating agent comprising at least one majorly abundant  
 PT extracellular product of pathogens, useful for promoting an effective  
 PT immune response in a mammalian host against infectious pathogens, e.g.  
 PT viruses, fungi, or bacteria.  
 PT  
 XX Claim 13; Page 31; 45pp; English.  
 PS  
 XX The invention describes a vaccinating agent for promoting an effective  
 CC immune response in a mammalian host against an infectious pathogen from  
 CC the genus Mycobacterium. The vaccinating agent or the combination vaccine  
 CC of majorly abundant extracellular products of pathogens is useful for  
 CC promoting an effective immune response in a mammalian host against  
 CC infectious pathogens, such as protozoa, viruses, fungi, or bacteria  
 CC particularly from the genus Mycobacterium, e.g. M. tuberculosis, M.  
 CC bovis, M. marinum, M. kansasii, M. aviumintracellulare, M. fortuitum, M.  
 CC chelonae, M. scrofulaceum, M. leprae, M. africanum, M. ulcerans, or M.  
 CC microti. This sequence represents an extended N-terminal peptide from one  
 CC of a group of majorly abundant extracellular products of Mycobacterium  
 CC tuberculosis  
 XX  
 XX Sequence 47 AA;  
 SQ  
 Query Match 91.7%; Score 22; DB 6; Length 47;  
 Best Local Similarity 80.0%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 40 GYTVE 44  
 RESULT 7  
 AAR63236  
 ID AAR63236 standard; protein; 68 AA.  
 XX  
 XX AAR63236;  
 AC  
 XX 25-MAR-2003 (revised)  
 DT  
 XX 06-JUL-1995 (first entry)  
 DT  
 XX Neural thread protein AD2-2 T7.  
 DE  
 XX Neural thread protein AD2-2 T7; Alzheimer's; neuroectodermal tumours;  
 KW malignant astrocytomas; glioblastomas.  
 KW  
 XX Rattus rattus.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 67. .68  
 FT /note= "corresponding codons GCA CAT CAC GAC ATT TTA TAA"  
 FT  
 XX WO9423756-A1.  
 PN  
 XX 27-OCT-1994.  
 PD  
 XX 20-APR-1994; 94WO-US004321.  
 XX  
 XX 20-APR-1993; 93US-00050559.  
 PR  
 XX (GEHO ) GEN HOSPITAL CORP.  
 XX PA

XX De La Monte SM, Wands JR;  
 PI WPI; 1994-341497/42.  
 DR N-PSDB; AAT277872.  
 XX  
 XX Detection of neural thread proteins - to detect sporadic and familial  
 PT Alzheimer's disease, neuroectodermal tumours, malignant astrocytomas and  
 PT glioblastomas (Eng).  
 XX  
 PS Example 4; Fig 16a; 158pp; English.  
 XX  
 CC AAQ77872 encodes AAR63236 the AD2-2 T7 neural thread protein (NTP). These  
 CC sequences were used in the development of an antibody dependent method,  
 CC for the detection of NTPs. This new method could be used to diagnose  
 CC Alzheimer's disease (differentiating between sporadic and familial),  
 CC neuroectodermal tumours, malignant astrocytomas and glioblastomas.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 68 AA;  
 Query Match 91.7%; Score 22; DB 2; Length 68;  
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 Db ||||  
 43 GYAVE 47  
 RESULT 8  
 AAR95915  
 ID AAR95915 standard; protein; 68 AA.  
 AC AAR95915;  
 XX  
 DT 14-NOV-1996 (first entry)  
 XX  
 DE AD 2-2 human neural thread protein clone (partial sequence).  
 XX  
 KW Neural thread protein; NTP; diagnosis; detection; Alzheimer's disease;  
 KW neuroectodermal tumour; malignant astrocytoma; monoclonal antibody;  
 KW binding fragment.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9615272-A1.  
 XX  
 PD 23-MAY-1996.  
 XX  
 PF 14-NOV-1995; 95WO-US017111.  
 XX  
 PR 14-NOV-1994; 94US-00340426.  
 XX  
 PA (GEO) GEN HOSPITAL CORP.  
 XX  
 PI De La Monte S, Wands JR;  
 XX  
 DR WPI; 1996-259865/26.  
 DR N-PSDB; AAT27754.  
 XX  
 XX Detection of neural thread protein in diagnosis of Alzheimer's disease -  
 PT also NTP DNA and protein sequences used in gene and anti:sense therapy.  
 PS  
 PS Example 4c; Fig 16A; 238pp; English.  
 XX  
 CC A method for detecting the presence of neural thread protein (NTP) having  
 CC a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human subject  
 CC comprises (a) contacting a sample from a human subject that is suspected  
 CC of containing the NTP with at least one molecule capable of binding to  
 CC the protein; and (b) detecting any of the molecule bound to the protein.  
 CC The binding molecule is selected from an antibody free of natural  
 CC impurities, a monoclonal antibody or a binding fragment of either of

CC these. The method may be used for diagnosing the presence of Alzheimer's  
 CC disease, neuroectodermal tumours and a malignant astrocytoma in a human.  
 CC A number of clones of neural thread protein were isolated from healthy 17  
 CC -18 week old fetal human brain (HB) 2 year old temporal lobe neocortex  
 CC and end stage Alzheimer's disease (AD) cerebral cortex. See AAT27753-75  
 XX  
 SQ Sequence 68 AA;  
 Query Match 91.7%; Score 22; DB 2; Length 68;  
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 Db ||||  
 43 GYAVE 47  
 RESULT 9  
 AAE29149  
 ID AAE29149 standard; protein; 68 AA.  
 XX  
 AC AAE29149;  
 XX  
 DT 27-JAN-2003 (first entry)  
 XX  
 DE Neural thread protein (NTP) #7.  
 XX  
 KW Neural thread protein; NTP; hyperplasia; hypertrophy; arteriosclerosis;  
 KW haemorrhoid; gene therapy; tumour; vascular disease; atherosclerosis;  
 KW inflammatory disease; nutritional deficiency disease; genetic disease;  
 KW autoimmune disease; metabolic disease; traumatic disease; intoxication;  
 KW infectious disease; congenital malformation; enzyme deficiency disease;  
 KW amyloid disease; fibrosis disease; storage disease; radiation disease;  
 KW poisoning; environmental disease; endocrine disease; protein therapy;  
 KW degenerative disease; mechanical disease.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200274323-A2.  
 XX  
 PD 26-SEP-2002.  
 XX  
 PF 08-MAR-2002; 2002WO-IB001959.  
 XX  
 PR 08-MAR-2001; 2001US-0273957P.  
 XX  
 PA (AVER/) AVERBACK P.  
 XX  
 PI Averbach P;  
 XX  
 DR WPI; 2002-759864/82.  
 XX  
 PT Treating a condition in a patient requiring removal or destruction of  
 PT cells, such as a benign or malignant tumor of a tissue or an inflammatory  
 PT disease, comprises administering a neural thread protein (NTP) or a NTP  
 PT gene to a mammal.  
 XX  
 PS Claim 23; Fig 8; 70pp; English.  
 XX  
 CC The invention relates to a method for treating a condition in a patient  
 CC requiring removal or destruction of cells. The method involves  
 CC administering to a mammal a neural thread protein (NTP), or administering  
 CC to a tumour or other target cell a NTP gene, where the expression of the  
 CC NTP gene is induced resulting in expression of the NTP protein. The  
 CC method and NTP are useful for treating a condition in a patient requiring  
 CC removal or destruction of cells, such as a benign or malignant tumour of  
 CC a tissue, a hyperplasia, hypertrophy, or overgrowth of a tissue,  
 CC preferably tonsillar hypertrophy or prostatic hyperplasia, a virally,  
 CC bacterially, or parasitically altered tissue, or a malformation of a  
 CC tissue. Other conditions include a cosmetic modification to a tissue,  
 CC such as removal of unwanted facial hair, warts or unwanted fatty tissue,  
 CC a vascular disease, particularly atherosclerosis or arteriosclerosis,  
 CC haemorrhoids, or varicose veins, an inflammatory disease, autoimmune

CC disease, metabolic disease, hereditary/genetic disease, traumatic disease  
 CC or physical injury, nutritional deficiency disease, infectious disease,  
 CC congenital malformation, amyloid disease, fibrosis disease, storage  
 CC disease, enzyme deficiency disease, poisoning, intoxication, degenerative  
 CC disease, radiation disease, environmental disease, endocrine disease or  
 CC mechanical disease. The invention is useful in protein therapy and gene  
 CC therapy. The present sequence is NTP protein  
 XX  
 SQ

Query Match 91.7%; Score 22; DB 5; Length 68;  
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 DB 43 GYAVE 47

RESULT 10  
 ID ABR63247 standard; protein; 68 AA.

AC ABR63247;  
 DT 28-AUG-2003 (first entry)

DE 68 amino acid neural thread protein.

KW Cytostatic; Antibacterial; Immunosuppressive; Antiinflammatory;  
 KW neural thread protein; NTP; tumour.

OS Unidentified.

XX WO2003008443-A2.

XX 30-JAN-2003.

XX 19-JUL-2002; 2002WO-CA001105.

XX 19-JUL-2001; 2001US-0306150P.

XX 19-JUL-2001; 2001US-0306161P.

XX 16-NOV-2001; 2001US-0331477P.

XX (NYMO-) NYMOX CORP.

XX Averbach PA;

XX WPI; 2003-247999/24.

XX Novel neural thread protein peptide, referred as cell death peptide,  
 XX useful for treating prostatic hyperplasia, psoriasis, eczema, dermatosis,  
 XX atherosclerosis, cosmetic modification to skin, throat, mouth, muscle.

XX Disclosure; Fig 8; 77pp; English.

XX The present invention relates to a neural thread protein (NTP) peptide  
 XX referred to as cell death peptide. Thought to be cytostatic,  
 XX antibacterial, immunosuppressive and antiinflammatory. It is useful for  
 XX treating a condition in a patient requiring removal or destruction of  
 XX cells, for treating a condition such as benign or malignant tumor,  
 XX inflammatory disease, autoimmune disease and infectious disease. The  
 XX peptide useful for treatment is derived from the amino acid sequence for  
 XX a pancreatic thread protein. The peptide is conjugated, linked or bound  
 XX to a molecule chosen from antibody or its fragment, antibody-like binding  
 XX molecule, where the molecule has a higher affinity for binding to a tumor  
 XX cell, other target than binding to other cells. Treatment using NTP peptides  
 XX can remove benign tumors with less risk and fewer of the undesirable side  
 XX effects of surgery. The present sequence is an NTP amino acid sequence

SQ Sequence 68 AA;

Query Match 91.7%; Score 22; DB 6; Length 68;

Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 DB 43 GYAVE 47

RESULT 11  
 ID ABU02979 standard; protein; 68 AA.

AC ABU02979;

XX 20-JAN-2003 (first entry)

XX Human neural thread protein AD7C-NTP, protein fragment #7.

KW Neural thread protein; NTP-peptide; AD7C-NTP; surgical excision;  
 KW transplantation; grafting; chemotherapy; immunotherapy; vaccination;  
 KW ablation; cryotherapy; laser therapy; phototherapy; gene therapy;  
 KW radiation; tumour; hyperplasia; hypertrophy; overgrowth of tissue;  
 KW malformation of tissue; tonsillary hypertrophy; prostatic hyperplasia;  
 KW cosmetic modification; vascular disease; atherosclerosis;  
 KW arteriosclerosis; haemorrhoid; varicose vein; inflammatory disease;  
 KW autoimmune disease; metabolic disease; traumatic disease;  
 KW physical injury; nutritional deficiency disease; infectious disease;  
 KW amyloid disease; fibrosis disease; storage disease;  
 KW congenital malformation; enzyme deficiency disease; poisoning;  
 KW intoxication; environmental disease; radiation disease;  
 KW endocrine disease; degenerative disease; mechanical disease.

OS Homo sapiens.

XX WO200297030-A2.

XX 05-DEC-2002.

XX 24-MAY-2002; 2002WO-CA000759.

XX 25-MAY-2001; 2001US-0293156P.

XX (NYMO-) NYMOX CORP.

XX Averbach PA;

XX WPI; 2003-041406/03.

XX Novel peptides similar in amino acid sequence to neural thread proteins  
 XX (NTP), useful for treating unwanted cellular proliferations such as  
 XX malignant tumors and prostatic hyperplasia.

XX Disclosure; Fig 8; 78pp; English.

XX The invention describes an NTP-peptide (I) comprising at least one amino  
 XX acid sequence corresponding to part of the amino acid sequence of a  
 XX neural thread protein, AD7C-NTP. The invention provides a method of  
 XX treating a condition requiring removal or destruction of cells of a  
 XX mammal comprising administering to a mammal, a therapeutic amount of (I).  
 XX The treatment is administered to the mammal before, during or after  
 XX surgical excision, transplantation, grafting, chemotherapy,  
 XX immunotherapy, vaccination, thermal or electrical ablation, cryotherapy,  
 XX laser therapy, phototherapy, gene therapy and/or radiation. The method is  
 XX useful for treatment of benign or malignant tumour; hyperplasia,  
 XX hypertrophy or overgrowth of tissue; virally, bacterially or  
 XX parasitically altered tissue; malformation of tissue selected from lung,  
 XX breast, stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney,  
 XX sinus, colon, intestine, rectum, esophagus, heart, spleen, salivary  
 XX gland, blood, brain and its coverings, spinal cord, muscle, connective  
 XX tissue, adrenal, parathyroid, thyroid, uterus, testis, pituitary,  
 XX reproductive organs, liver, hair, gall bladder, eye, ear, nose, throat,  
 XX tonsils, mouth and lymph nodes and lymphoid system; tonsillary  
 XX hypertrophy; prostatic hyperplasia; cosmetic modification to a tissue;



CC vascular disease (atherosclerosis or arteriosclerosis); haemorrhoids;  
CC varicose veins; inflammatory disease; autoimmune disease; metabolic  
CC disease; hereditary/genetic disease; traumatic disease; physical injury;  
CC nutritional deficiency disease; infectious disease; amyloid disease;  
CC fibrosis disease; storage disease; congenital malformation; enzyme  
CC deficiency disease; poisoning; intoxication; environmental disease;  
CC radiation disease; endocrine disease; degenerative disease and mechanical  
CC disease. This is the amino acid sequence of a human neural thread protein  
XX AD7C-NTP protein fragment  
XX  
SQ Sequence 68 AA;

Query Match 91.7%; Score 22; DB 6; Length 68;  
Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 43 GYAVE 47

## RESULT 12

ID ABP59929 standard; protein; 68 AA.

XX ABP59929;

AC ABP59929;

DT 08-SEP-2003 (first entry)

DE Human 68 amino acid neural thread protein.

XX Human; tumour; cancer; neural thread protein; NTP; cell removal;  
KW cell destruction; antiporiatic; antimicrobial; immunosuppressive;  
KW antiinflammatory; dermatological; antiarteriosclerotic; vasotropic;  
KW gene therapy.

XX Homo sapiens.

XX WO2003044053-A2.

XX 30-MAY-2003.

XX 18-NOV-2002; 2002WO-CA001757.

XX 16-NOV-2001; 2001US-0331477P.

XX (NYMO-) NYMOX CORP.

XX Averbach P, Gemmell J;

XX WPI; 2003-457592/43.

XX New neural thread protein (NTP), useful for preparing a composition for  
PT treating or preventing a condition in a mammal requiring removal or  
PT destruction of cells, e.g. psoriasis, eczema, atherosclerosis or  
PT inflammatory disease.

XX Disclosure; Fig 6; 98pp; English.

XX The present invention relates to peptides derived from the human neural  
CC thread protein (NTP). The peptides are useful for preparing a composition  
CC for treating or preventing a condition in a mammal requiring removal or  
CC destruction of cells, comprising tonsillary hypertrophy, prostatic  
CC hyperplasia, psoriasis, eczema, dermatosis, cosmetic modification to a  
CC breast, connective, skin, eye, ear, nose, throat, mouth or muscle tissue,  
CC varicose veins, atherosclerosis, inflammatory, metabolic, infectious,  
CC fibrosis, endocrine or autoimmune disease, or stenosis, restenosis,  
CC occlusion or blockage of an artery or of a stent placed or implanted in  
CC an artery. The present sequence is an NTP protein used to produce  
CC peptides of the invention

XX Sequence 68 AA;

Query Match 91.7%; Score 22; DB 6; Length 68;  
Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 43 GYAVE 47

## RESULT 13

AAE33196  
ID AAE33196 standard; protein; 68 AA.

XX AAE33196;

AC AAE33196;

DT 16-APR-2003 (first entry)

DE Neural thread protein (NTP) #7.

XX Cell death; tissue necrosis; neural thread protein; NTP; amyloidosis;  
KW stroke; brain tumour; Pick's disease; Parkinson's disease; glioma;  
KW Alzheimer's disease; gene therapy.

XX Unidentified.

XX WO200289841-A2.

XX 14-NOV-2002.

XX 06-MAY-2002; 2002WO-CA000681.

XX 04-MAY-2001; 2001US-0288463P.

XX (NYMO-) NYMOX CORP.

XX Averbach PA;

XX WPI; 2003-120506/11.

XX Preventing, controlling, modulating, ameliorating and/or treating cell  
PT death or tissue necrosis using antibodies to neural thread proteins,  
PT useful in disorders such as stroke, brain tumor, glioma and Alzheimer's  
PT disease.

XX Disclosure; Fig 11; 60pp; English.

XX The invention relates to a method of preventing, and/or inhibiting cell  
CC death and/or tissue necrosis in live tissue containing neural thread  
CC proteins (NTP). The method involves contacting the live tissue with at  
CC least one antibody, fragment or derivative that recognises NTP, where the  
CC antibody, fragment or derivative is present to prevent, control,  
CC ameliorate and/or inhibit cell death and/or tissue necrosis caused by the  
CC presence of NTP. Methods and compositions of the invention are useful for  
CC preventing, modulating, controlling and/or treating disorders associated  
CC with cell death and/or tissue necrosis such as stroke, brain tumour,  
CC Pick's disease, Parkinson's disease, amyloidosis, glioma and  
CC Alzheimer's disease. The invention is useful in gene therapy. The present  
CC sequence is NTP protein

XX Sequence 68 AA;

Query Match 91.7%; Score 22; DB 6; Length 68;  
Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 43 GYAVE 47

## RESULT 14

ABJ19452  
ID ABJ19452 standard; protein; 68 AA.

XX ABJ19452;  
 AC 27-MAR-2003 (first entry)  
 DT  
 DE 68-mer neural thread protein.  
 DE  
 KW Nootropic; neuroprotective; cell death; tissue necrosis; NTP;  
 KW neural thread protein; neurodegenerative disorder; Alzheimer's disease.  
 XX  
 OS Unidentified.  
 XX WO200292115-A2.  
 PN  
 XX 21-NOV-2002.  
 PD  
 XX 16-MAY-2002; 2002WO-CA000712.  
 PF  
 XX 16-MAY-2001; 2001US-0290971P.  
 PR  
 XX (NYMO-) NYMOX CORP.  
 PA  
 XX Averbach PA;  
 PI  
 XX WPI; 2003-129234/12.  
 DR  
 XX Preventing and/or inhibiting cell death and/or tissue necrosis in a  
 XX tissue for treating a neurodegenerative disorder, e.g. Alzheimer's  
 PT disease, by contacting the live tissue with at least one segment of  
 PT neural thread proteins (NTP).  
 PT  
 PS Disclosure; Fig 8; 60pp; English.  
 PS  
 XX The invention relates to a novel method for preventing and/or inhibiting  
 CC cell death and/or tissue necrosis in a tissue comprising contacting the  
 CC live tissue with at least one segment of neural thread proteins (NTP).  
 CC The methods are composition are useful for treating a neurodegenerative  
 CC disorder, such as Alzheimer's disease. This sequence represents an NTP  
 CC protein of the invention  
 XX  
 SQ Sequence 68 AA;  
 Query Match 91.7%; Score 22; DB 6; Length 68;  
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 43 GYAVE 47  
 RESULT 15  
 ADB37526  
 ID ADB37526 standard; protein; 68 AA.  
 AC  
 XX ADB37526;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Neural thread protein #5.  
 XX  
 KW Cytostatic; Antitumour; Antipsoriatic; Dermatological;  
 KW Antiatherosclerotic; Antiarteriosclerotic; Vasotropic; Antiinflammatory;  
 KW Immunosuppressive; Tranquillizer; Antiemetic; Virucide; AD7c-NTP;  
 KW neural thread protein; neuritic sprouting.  
 XX  
 OS Unidentified.  
 XX  
 XX WO2003008444-A2.  
 PN  
 XX 30-JAN-2003.  
 PD  
 XX 19-JUL-2002; 2002WO-CA001106.  
 PF

XX 19-JUL-2001; 2001US-0306150P.  
 PR 19-JUL-2001; 2001US-0306161P.  
 PR 16-NOV-2001; 2001US-0331477P.  
 XX  
 PA (NYMO-) NYMOX CORP.  
 XX  
 XX Averbach PA, Gemmell J;  
 PI  
 XX WPI; 2003-248000/24.  
 DR  
 XX Novel Related peptide or AD7c-neural thread peptide, useful for treating  
 PT unwanted cellular proliferations, glandular hyperplasia, unwanted facial  
 PT hair, warts and unwanted fatty tissue.  
 PT  
 XX Disclosure; Fig 8; 109pp; English.  
 PS  
 XX The present invention relates to AD7c-neural thread protein (NTP) and  
 CC related proteins and peptides (1; ADB37528-ADB37641). The sequences are  
 CC useful for treating a condition in a patient requiring removal or  
 CC destruction of cells. The condition can be selected from benign or  
 CC malignant tumour of a tissue, hyperplasia, hypertrophy or overgrowth of a  
 CC tissue, virally, bacterially or parasitically altered tissue, or  
 CC malformation of a tissue, where the tissue is selected from lung, breast,  
 CC stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney, sinus,  
 CC colon, intestine, stomach, rectum, oesophagus, heart, spleen, salivary  
 CC gland, blood, brain and its coverings, spinal cord and its coverings,  
 CC muscle, connective tissue, adrenal, parathyroid, thyroid, uterus, testis,  
 CC pituitary, reproductive organs, liver, gall bladder, eye, ear, nose,  
 CC throat, tonsils, mouth, lymph nodes and lymphoid tissue. The condition is  
 CC preferably tonsillary hypertrophy, prostatic hyperplasia, psoriasis,  
 CC eczema, dermatosis, cosmetic modification to a tissue (skin, eye, ear,  
 CC nose, throat, mouth, muscle, connective, hair or breast tissue), vascular  
 CC disease (atherosclerosis or arteriosclerosis), haemorrhoids, varicose  
 CC veins, inflammatory disease, autoimmune disease, metabolic disease,  
 CC hereditary/genetic disease, traumatic disease or physical injury,  
 CC nutritional deficiency disease, infectious disease, amyloid disease,  
 CC fibrosis disease, storage disease, congenital malformation, enzyme  
 CC deficiency disease, poisoning, intoxication, environmental disease,  
 CC radiation disease, endocrine disease, degenerative disease and mechanical  
 CC disease. The peptides are useful for treating unwanted cellular  
 CC proliferations, glandular (e.g. prostate) hyperplasia, unwanted facial  
 CC hair, warts and unwanted fatty tissue, or for preparing antibodies that  
 CC recognize and/or bind to Related proteins, Related peptides or NTP  
 CC peptides. The present sequence was used to illustrate the invention.  
 XX  
 SQ Sequence 68 AA;  
 Query Match 91.7%; Score 22; DB 7; Length 68;  
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 43 GYAVE 47  
 RESULT 16  
 ADL96027  
 ID ADL96027 standard; protein; 68 AA.  
 XX  
 AC ADL96027;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Human neural thread protein, NTP68.  
 XX  
 KW Human; neural thread protein; NTP122; NTP112; NTP106; NTP98; NTP75;  
 KW NTP68; NTP61; stenosis; stent; tumour; prostatic hyperplasia; psoriasis;  
 KW eczema; haemorrhoid; atherosclerosis; inflammatory disease;  
 KW autoimmune disease; metabolic disease; hereditary disease;  
 KW genetic disease; traumatic disease; physical injury;  
 KW nutritional deficiency disease; infectious disease; amyloid disease;

KW Alzheimer's disease; storage disease; congenital malformation;  
 KW enzyme deficiency disease; poisoning; intoxication;  
 KW environmental disease; radiation disease; endocrine disease;  
 KW degenerative disease; mechanical disease.

OS Homo sapiens.

PN US2003166569-A1.

XX 04-SEP-2003.

XX 15-NOV-2002; 2002US-00294891.

XX 16-NOV-2001; 2001US-0331477P.

XX (AVER/) AVERBACK P.

PA (GEMM/) GEMMELL J.

PI Averbach P, Gemmell J;

XX WPI; 2003-898099/82.

DR New neural thread protein or its variants, useful for treating tumors and  
 XX other conditions requiring the removal or destruction of cells (e.g.  
 PT prostatic hyperplasia, psoriasis, eczema, hemorrhoids or  
 PT atherosclerosis).

PS Disclosure; SEQ ID NO 6; 32pp; English.

CC The invention relates to a peptide, or its homologue, derivative,  
 CC fragment, variant or mimetic, comprising at least one neural thread  
 CC protein (NTP) peptide appearing as ADU96029-ADU96069, derived from  
 CC NTP122, 112, 106, 98, 75, 68 or 61. Also included are a nucleic acid  
 CC encoding an amino acid sequence corresponding to the above peptide, a  
 CC composition comprising one or more peptides or nucleic acids cited above  
 CC and a carrier, a method of treating a condition in a mammal requiring  
 CC removal or destruction of cells (comprising administering to the mammal  
 CC an amount of the peptide cited above) and a method of preventing or  
 CC inhibiting the stenosis, occlusion or blockage of a stent, comprising  
 CC coating the stent with an amount of the above peptide. The peptide  
 CC further comprises an amino acid in a reverse-D order based on the above  
 CC amino acid sequences. The composition and methods are useful in treating  
 CC tumors and other conditions requiring the removal or destruction of  
 CC cells (e.g. prostatic hyperplasia, psoriasis, eczema, haemorrhoids or  
 CC atherosclerosis). These may also be used in treating inflammatory  
 CC diseases, autoimmune diseases, metabolic diseases, hereditary/genetic  
 CC diseases, traumatic diseases or physical injuries, nutritional deficiency  
 CC diseases, infectious diseases, amyloid diseases e.g. Alzheimer's disease,  
 CC storage diseases, congenital malformation, enzyme deficiency diseases,  
 CC poisoning, intoxication, environmental diseases, radiation diseases,  
 CC endocrine diseases, degenerative diseases or mechanical diseases. The  
 CC present sequence is a human NTP protein from which the peptides of the  
 CC invention are derived.

XX Sequence 68 AA;

Query Match 91.7%; Score 22; DB 7; Length 68;

Best Local Similarity 80.0%; Pred. No. 3.8e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 43 GYAVE 47

RESULT 17

AAW02576

ID AAW02576 standard; protein; 78 AA.

XX AAW02576;

XX 24-JAN-1997 (first entry)

XX

DE E. coli plasmid R67 dihydrofolate reductase.

XX Hypermutation; reverse transcription; biasing dNTP concentration;  
 KW Escherichia coli; dihydrofolate reductase; DHFR; plasmid R67;  
 KW polymerase chain reaction; PCR; amplification; primer.

OS Escherichia coli.

XX WO9617056-A1.

PN 06-JUN-1996.

XX 01-DEC-1995; 95WO-EP004749.

XX 02-DEC-1994; 94EP-00402774.

PR 19-MAY-1995; 95US-00447173.

XX (INSP ) INST PASTEUR.

XX Wain-Hobson S;

PI WPI; 1996-277776/28.

DR N-PSDB; AAT32663.

XX Inducing hyper-mutation(s) in DNA or RNA - by transcribing RNA into DNA  
 PT in the presence of varying biased concns. of deoxy:nucleotide  
 PT tri:phosphate(s).

PS Example 2; Fig 5; 72pp; English.

XX Hypermutations can be introduced into RNA or DNA by reverse transcription  
 CC (RT) in the presence of a biasing concn. of dNTP. In an example, a  
 CC dihydrofolate reductase (DHFR) gene encoded by the E. coli plasmid R67  
 CC was utilised as the target DNA sequence. The DHFR gene was 1st amplified  
 CC from the plasmid using PCR and the primer pair AAT32664/65. The prod.  
 CC then underwent RT in a biasing concn. of dNTP, and the resulting cDNA  
 CC amplified using the above primer pair. The prod. was cleaved using the  
 CC introduced restriction sites, and ligated into the pRC39A expression  
 CC vector for the transformation of E. coli cells. The cells were then  
 CC plated out on to standard Lauria broth supplemented with trimethoprim  
 CC (TMP) and ampicillin. To identify the nature of the hypermutated R67  
 CC genes, TMP resistant colonies were grown up singly, plasmid DNA extracted  
 CC and the DNA sequenced and compared to the wild type DHFR DNA sequence  
 CC AAT32663, which encodes AAW02576

XX Sequence 78 AA;

Query Match 91.7%; Score 22; DB 2; Length 78;

Best Local Similarity 80.0%; Pred. No. 4.4e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 54 GYAVE 58

RESULT 18

AAAR95919

ID AAR95919 standard; protein; 78 AA.

XX AAR95919;

XX 24-JAN-1997 (first entry)

DE E. coli plasmid R67 dihydrofolate reductase.

XX Hypermutation; reverse transcription; biasing dNTP concentration;  
 KW Escherichia coli; dihydrofolate reductase; DHFR; plasmid R67;  
 KW polymerase chain reaction; PCR; amplification; primer.

XX Escherichia coli.

XX EP714980-A1.

XX PD 05-JUN-1996.  
 XX PF 02-DEC-1994; 94EP-00402774.  
 XX PR 02-DEC-1994; 94EP-00402774.  
 XX PA (INSP ) INST PASTEUR.  
 XX FI Martinez MA, Main-Hobson S;  
 XX FI WPI; 1996-269713/28.  
 XX DR N-PSDB; AAT27833.  
 XX XX  
 PT Introducing hyper:mutation(s) into RNA and DNA by reverse transcription -  
 PT in presence of biasing concn. of dNTP, also hyper:mutated nucleic acid  
 PT and derived mutant proteins.  
 XX Example 2; Fig 5; 47pp; English.  
 XX Hypermutations can be introduced into RNA or DNA by reverse transcription  
 CC (RT) in the presence of a biasing concn. of dNTP. In an example, a  
 CC dihydrofolate reductase (DHFR) gene encoded by the E. coli plasmid R67  
 CC was utilised as the target DNA sequence. The DHFR gene was 1st amplified  
 CC from the plasmid using PCR and the primer pair AAT27831/32. The prod.  
 CC then underwent RT in a biasing concn. of dNTP, and the resulting cDNA  
 CC amplified using the above primer pair. The prod. was cleaved using the  
 CC introduced restriction sites, and ligated into the pRC99A expression  
 CC vector for the transformation of E. coli cells. The cells were then  
 CC plated out on to standard Lauria broth supplemented with trimethoprim  
 CC (TMP) and ampicillin. To identify the nature of the hypermutated R67  
 CC genes, TMP resistant colonies were grown up singly, plasmid DNA extracted  
 CC and the DNA sequenced and compared to the wild type DHFR DNA sequence  
 CC AAT27833, which encodes AAR95919  
 XX AAT27833, which encodes AAR95919  
 XX SQ Sequence 78 AA;  
 Query Match 91.7%; Score 22; DB 2; Length 78;  
 Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 Db 54 GYAVE 58  
 |||||  
 RESULT 19  
 AAY60563  
 ID AAY60563 standard; protein; 83 AA.  
 XX AC AAY60563;  
 XX DT 31-JAN-2000 (first entry)  
 XX DE Human normal bladder tissue EST encoded protein 235.  
 XX KW Human; bladder; treatment; EST; expressed sequence tag; cytostatic;  
 XX KW Cancer; gene therapy.  
 XX OS Homo sapiens.  
 XX PN DE19818620-A1.  
 XX PD 28-OCT-1999.  
 XX PF 21-APR-1998; 98DE-01018620.  
 XX PR 21-APR-1998; 98DE-01018620.  
 XX PA (META-) METAGEN GFS GENOMFORSCHUNG MBH.  
 XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 XX

DR WPI; 1999-602416/52.  
 XX N-PSDB; AAZ42237.  
 PT New polypeptides and their nucleic acids, useful for treatment of bladder  
 PT tumor and identification of therapeutic agents.  
 XX Claim 23; Page 340; 366pp; German.  
 XX This invention describes novel polypeptide fragment sequences (I) and  
 CC their encoding nucleic acids (II) which are highly expressed in normal  
 CC bladder tissue and have cytostatic activity. (II) are used for  
 CC recombinant expression of (I) and to isolate complete genes. (I) are used  
 CC to identify agents suitable for the treatment of bladder tumours, to  
 CC directly treat this form of cancer (including expression from gene  
 CC therapy vectors), or are used in a preparation for cancer treatment. (I)  
 CC is also used for the generation of specific antibodies. (II) are  
 CC identified by assembling ESTs (expressed sequence tags) from a particular  
 CC tissue type before comparison of expression patterns. This allows a  
 CC significantly longer fragment of the gene to be revealed, and therefore  
 CC reduces the number of failures because of ESTs from different libraries  
 CC representing different parts of the same unknown gene distorting the  
 CC estimated frequency of occurrence in a particular tissue. AAY60329-Y60591  
 CC represent protein fragments encoded by the human normal bladder tissue  
 CC cDNA library derived EST fragments represented in AAZ42122-Z42248  
 XX cDNA library derived EST fragments represented in AAZ42122-Z42248  
 XX SQ Sequence 83 AA;  
 Query Match 91.7%; Score 22; DB 2; Length 83;  
 Best Local Similarity 80.0%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 Db 14 GYTV 18  
 |||||  
 RESULT 20  
 AAR91284  
 ID AAR91284 standard; protein; 104 AA.  
 XX AC AAR91284;  
 XX DT 01-APR-1997 (first entry)  
 XX DE LERK-6 exon polypeptide.  
 XX KW LERK-6; hek; elk; cell surface receptors; culture; reagent; neurons;  
 XX KW disorder; injury; delivery agent; diagnostic; therapeutic.  
 XX OS Homo sapiens.  
 XX PN WO9610911-A1.  
 XX PD 18-APR-1996.  
 XX PF 04-OCT-1995; 95WO-US012779.  
 XX PR 05-OCT-1994; 94US-00318393.  
 XX PR 03-OCT-1995; 95US-00538709.  
 XX PA (IMMV ) IMMUNEX CORP.  
 XX FI Cerretti DP;  
 XX DR WPI; 1996-209575/21.  
 XX DR N-PSDB; AAT14010.  
 XX Isolated DNA encoding cytokine designated LERK-6 which binds to hek and  
 PT elk cell surface receptors - useful for drug delivery and screening  
 PT procedures.  
 XX Claim 3; Page 36; 44pp; English.  
 XX

CC The LERK-6 polypeptide can be used to isolate cells expressing hek/ek  
 CC cell surface receptors, or to measure the biological activity of such  
 CC receptors. The protein may also be used as a delivery agent, taking  
 CC diagnostic and therapeutic agents to cells expressing such receptors.  
 CC LERK-6 can also be used as a tissue culture reagent to enhance the  
 CC viability or prolong the lifespan of the neurons. Neural tissue disorders  
 CC and injuries may be treated by contact with the polypeptide  
 XX  
 SQ Sequence 104 AA;

Query Match 91.7%; Score 22; DB 2; Length 104;  
 Best Local Similarity 80.0%; Pred. No. 6e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 11 GYTV 15

RESULT 21  
 ID AAY06821 standard; protein; 104 AA.  
 XX AC AAY06821;  
 XX DT 24-JUN-1999 (first entry)  
 XX DE Amino acid sequence encoded by human LERK-6 exon.  
 XX KW LERK-6 polypeptide; hek receptor; elk receptor; human; murine; injury;  
 XX KW cell proliferation; neural growth; neural tissue; neurological disease;  
 XX KW neurodegenerative; excitotoxicity.  
 XX OS Homo sapiens.  
 XX PN WO9910495-A1.  
 XX PD 04-MAR-1999.  
 XX PF 27-AUG-1998; 98WO-US017772.  
 XX PR 29-AUG-1997; 97US-00920440.  
 XX FA (IMMV ) IMMUNEX CORP.  
 XX PI Cerretti DP;  
 XX PS Claim 6; Page 24; 46pp; English.  
 XX CC The invention relates murine and human LERK-6 polypeptides that bind to  
 CC hek/elk receptors. Host cells transfected or transformed with vectors  
 CC comprising the LERK-6 nucleic acid sequences are used for the recombinant  
 CC production of the proteins. LERK-6 polypeptides may be useful in the  
 CC enhancement, stimulation, proliferation or growth of cells expressing the  
 CC hek or elk receptor. The ligand and receptor complex may be involved in  
 CC neural growth, development and/or maintenance. LERK-6 can be used for  
 CC treating disorders of neural tissue such as injury or neurological  
 CC diseases, either chronic or acute. LERK-6 may be employed in treating  
 CC neurodegenerative conditions where there is neural death, excitotoxicity.  
 CC In addition, they may be administered to a mammal to exert a trophic  
 CC effect on neural tissue. They can also be used as reagents for those  
 CC conducting quality assurance studies e.g. to monitor shelf life and  
 CC stability of elk protein under different conditions. The polypeptides can  
 CC also be used as carriers for delivering agents attached to cells bearing  
 CC the elk or hek cell surface receptor. The present sequence represents a  
 CC sequence encoded by an exon from human LERK-6 cDNA  
 XX  
 SQ Sequence 104 AA;

Query Match 91.7%; Score 22; DB 2; Length 104;  
 Best Local Similarity 80.0%; Pred. No. 6e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 11 GYTV 15

RESULT 22  
 ID ADH86927 standard; protein; 113 AA.  
 XX AC ADH86927;  
 XX DT 22-APR-2004 (first entry)  
 XX DE Enterococcus faecalis polypeptide #1407.  
 XX KW Enterococcus faecalis infection; transcription regulatory element;  
 XX KW antibacterial.  
 XX OS Enterococcus faecalis.  
 XX PN US6617156-B1.  
 XX PD 09-SEP-2003.  
 XX PF 13-AUG-1998; 98US-00134000.  
 XX PR 15-AUG-1997; 97US-0055778P.  
 XX PA (DOUC/) DOUCETTE-STAMM L A.  
 XX PA (BUSH/) BUSH D.  
 XX PI Doucette-Stamm LA, Bush D;  
 XX DR WPI; 2003-895394/82.  
 XX DR N-PSDB; ADH83522.  
 XX PT New nucleic acid comprising a sequence encoding an Enterococcus faecalis  
 XX PT polypeptide, useful for preparing a composition for diagnosing or  
 XX PS treating E. faecalis infection.  
 XX PS Disclosure; SEQ ID NO 4812; 193pp; English.  
 XX CC The invention relates to Enterococcus faecalis polynucleotides and  
 XX CC polypeptides. The invention also relates to a recombinant expression  
 XX CC vector comprising a polynucleotide operably linked to a transcription  
 XX CC regulatory element, a cell comprising a recombinant vector, a method for  
 XX CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising  
 XX CC a sequence not given in the specification, a recombinant vector  
 XX CC comprising the nucleic acid and a cell comprising the recombinant vector.  
 XX CC The polynucleotides can be used to detect the presence of E. faecalis in  
 XX CC a sample. The sequences are useful for preparing a composition for  
 XX CC diagnosing or treating Enterococcus faecalis infection. This sequence  
 XX CC represents an E. faecalis polypeptide of the invention.  
 XX  
 SQ Sequence 113 AA;

Query Match 91.7%; Score 22; DB 7; Length 113;  
 Best Local Similarity 80.0%; Pred. No. 6.5e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 6 GYSVE 10

RESULT 23  
 ID AAG30135 standard; protein; 125 AA.

XX AC AAG30135;  
 XX DT 17-OCT-2000 (first entry)  
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 35973.  
 XX KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hydrihydation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX OS Arabidopsis thaliana.  
 XX EN EP1033405-A2.  
 XX PD 06-SEP-2000.  
 XX PF 25-FEB-2000; 2000EP-00301439.  
 XX PR 25-FEB-1999; 99US-0121835P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 30-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132407P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 11-MAY-1999; 99US-0132863P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 21-MAY-1999; 99US-0135124P.  
 PR 24-MAY-1999; 99US-0135353P.  
 PR 25-MAY-1999; 99US-0135629P.  
 PR 27-MAY-1999; 99US-0136021P.  
 PR 28-MAY-1999; 99US-0136392P.  
 PR 01-JUN-1999; 99US-0136782P.  
 PR 03-JUN-1999; 99US-0137222P.  
 PR 04-JUN-1999; 99US-0137528P.  
 PR 07-JUN-1999; 99US-0137503P.  
 PR 08-JUN-1999; 99US-0137724P.  
 PR 10-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 14-JUN-1999; 99US-0138847P.  
 PR 16-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 17-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 23-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 23-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.

```
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.7%; Score 22; DB 3; Length 125;
Best Local Similarity 80.0%; Pred. No. 7.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 8 GYSVE 12

RESULT 24
ADH86269
ID ADH86269 standard; protein; 140 AA.
XX AC
XX ADH86269;
XX DT 22-APR-2004 (first entry)
XX XX
DE Enterococcus faecalis polypeptide #749.
XX XX
```

```
KW Enterococcus faecalis infection; transcription regulatory element;
KW antibacterial.
XX OS
XX Enterococcus faecalis.
PN US6617156-B1.
XX PD 09-SEP-2003.
XX PF 13-AUG-1998; 98US-00134000.
XX PR 15-AUG-1997; 97US-0055778P.
XX PA (DOUC/) DOUCETTE-STAMM L A.
XX PA (BUSH/) BUSH D.
XX PI Doucette-Stamm LA, Bush D;
XX DR WPI; 2003-895394/82.
XX DR N-PSDB; ADH82864.
XX PT New nucleic acid comprising a sequence encoding an Enterococcus faecalis
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating E. faecalis infection.
XX PS Disclosure; SEQ ID NO 4154; 193pp; English.
XX CC The invention relates to Enterococcus faecalis polynucleotides and
CC polypeptides. The invention also relates to a recombinant expression
CC vector comprising a polynucleotide operably linked to a transcription
CC regulatory element, a cell comprising a recombinant vector, a method for
CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising
CC a sequence not given in the specification, a recombinant vector
CC comprising the nucleic acid and a cell comprising the recombinant vector.
CC The polynucleotides can be used to detect the presence of E. faecalis in
CC a sample. The sequences are useful for preparing a composition for
CC diagnosing or treating Enterococcus faecalis infection. This sequence
CC represents an E. faecalis polypeptide of the invention.
XX SQ Sequence 140 AA;

Query Match 91.7%; Score 22; DB 7; Length 140;
Best Local Similarity 80.0%; Pred. No. 8.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 111 GYVE 115

RESULT 25
ABO84239
ID ABO84239 standard; protein; 143 AA.
XX AC ABO84239;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #16414.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX XX
```

PA (GENO-) GENOME THERAPEUTICS CORP.  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
PI WPI; 2003-615309/58.  
XX N-PSDB; ABD17810.  
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX Disclosure; SEQ ID NO 32985; 455pp; English.  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences AB067826-  
CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USFTO at  
CC seqdata.uspto.gov/sequence.html  
XX SQ Sequence 143 AA;  
Query Match 91.7%; Score 22; DB 7; Length 143;  
Best Local Similarity 80.0%; Pred. No. 8.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
DB 10 GYAVE 14  
|||  
RESULT 26  
AAE15402  
ID AAE15402 standard; protein; 147 AA.  
XX AAE15402;  
XX 26-MAR-2002 (first entry)  
XX HPV 16 E2 protein.  
XX Human papillomavirus;  
XX cancer; intraepithelial neoplasia; IN; squamous cell carcinoma; tumour;  
XX gene therapy; keratinocyte.  
XX Human papillomavirus.  
XX WO200187350-A2.  
XX 22-NOV-2001.  
XX 11-MAY-2001; 2001WO-US015407.  
XX 12-MAY-2000; 2000US-0203709P.  
XX (REGC ) UNIV CALIFORNIA.  
XX Sethi N, Palefsky J;  
XX WPI; 2002-082947/11.  
XX Novel gene therapy approach to specifically eliminate keratinocytes or  
PT other cells expressing human papilloma virus, involves transfecting cell

PT with a construct encoding HPV specific promoter induced by a HPV protein.  
XX Example 1; Fig 13; 72pp; English.  
XX The invention provides a novel gene therapy approach to specifically  
CC eliminate keratinocytes or other cells expressing early human  
CC papillomavirus (HPV) and which is minimally toxic to HPV-negative cells.  
CC The method involves transfecting a mammalian cell with a nucleic acid  
CC construct encoding a HPV specific promoter that is induced by a HPV  
CC protein where the promoter is operably linked to a nucleic acid  
CC comprising a cytotoxic gene such that the cell, when infected with a HPV,  
CC induces expression of the cytotoxic gene resulting in death of the  
CC mammalian cell. The method is useful for selectively killing a cancer  
CC cell comprising intraepithelial neoplasia (IN), anogenital cancer or a  
CC metastatic cell or solid tumour bearing a HPV. The nucleic acid construct  
CC comprising HPV promoter operably linked to a reporter gene is useful for  
CC selectively labelling a cell bearing HPV. The method is useful for  
CC specifically eliminating keratinocytes, particularly cells of anogenital  
CC squamous cell carcinoma (SCC), or other cells expressing early HPV e.g.  
CC HPV 16 genes and a cell comprising a wart. The present sequence is HPV 16  
CC E2 protein  
XX SQ Sequence 147 AA;  
Query Match 91.7%; Score 22; DB 5; Length 147;  
Best Local Similarity 80.0%; Pred. No. 8.7e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
DB 50 GYTVE 54  
|||  
RESULT 27  
ABB48719  
ID ABB48719 standard; protein; 148 AA.  
XX ABB48719;  
XX 05-FEB-2002 (first entry)  
XX Listeria monocytogenes protein #1423.  
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
XX vitamin B12; bacterial infection; disease.  
XX Listeria monocytogenes.  
XX WO200177335-A2.  
XX 18-OCT-2001.  
XX 11-APR-2001; 2001WO-FR001118.  
XX 11-APR-2000; 2000FR-00004629.  
XX (INSP ) INST PASTEUR.  
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Faihi H, Dehoux P;  
XX Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;  
XX Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;  
XX Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
XX Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;  
XX Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
XX Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;  
XX Rose M, Voss H;  
XX WPI; 2002-010914/01.  
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment  
PT and prevention of Listeria and related bacterial infections, and related  
PT polypeptides.  
XX



PS Claim 6; SEQ ID NO 1424; 192pp; French.

CC The present invention relates to the genome sequence of *Listeria*

CC monocytes EGD-e (see ABA03041). The genome sequence and fragments of

CC it are useful for selecting probes and primers for detecting genes in *L.*

CC monocytes and related organisms, and for studying genetic

CC polymorphisms and other genomes. The present sequence is a protein

CC encoded by the genome sequence of the present invention. Proteins

CC expressed from the genome sequence are useful for raising specific

CC antibodies, identification of *L. monocytogenes* and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin

CC B12. The genome sequence and proteins encoded by it are also useful for

CC selecting compounds that regulate gene expression and cell replication

CC and modulate *L. monocytogenes*-related diseases. In addition, the genome

CC sequence and proteins encoded by it are useful in pharmaceutical and

CC vaccine compositions for the treatment or prevention of infections by *L.*

CC monocytes and related organisms. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX

XX Sequence 148 AA;

Query Match 91.7%; Score 22; DB 5; Length 148;

Best Local Similarity 80.0%; Pred. No. 8.7e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 34 GYAVE 38

RESULT 28

ABU32695

ID ABU32695 standard; protein; 148 AA.

XX

AC ABU32695;

XX

XX 19-JUN-2003 (first entry)

XX

XX Protein encoded by Prokaryotic essential gene #18222.

XX

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

XX *Listeria* monocytes.

XX

XX WO200277183-A2.

XX

XX 03-OCT-2002.

XX

XX 21-MAR-2002; 2002WO-US009107.

XX

XX 21-MAR-2001; 2001US-00815242.

XX

XX 06-SEP-2001; 2001US-00948993.

XX

XX 25-OCT-2001; 2001US-0342923P.

XX

XX 08-FEB-2002; 2002US-00072851.

XX

XX 06-MAR-2002; 2002US-0362699P.

XX

XX (ELIT-) ELITRA PHARM INC.

XX

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

XX WPI; 2003-029926/02.

XX

XX N-PSDB; ACA36565.

XX

XX New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX

XX Claim 25; SEQ ID NO 60619; 176pp; English.

XX

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway;

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX

XX Sequence 148 AA;

Query Match 91.7%; Score 22; DB 6; Length 148;

Best Local Similarity 80.0%; Pred. No. 8.7e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 34 GYAVE 38

RESULT 29

AAB99858

ID AAB99858 standard; protein; 153 AA.

XX

AC AAB99858;

XX

XX 21-SEP-2001 (first entry)

XX

XX *Physcomitrella* patens 26\_ppprot140\_E07rev protein.

XX

XX Tocopherol and carotenoid metabolism related protein; TCWRP; synthesis;

XX *Physcomitrella* patens; moss; algae; microorganism; fungus; plant;

XX identification; genome mapping; modulation; evolutionary study;

XX cellular production; fine chemical.

XX

XX *Physcomitrella* patens.

XX

XX WO200144276-A2.

XX

XX 21-JUN-2001.

XX

XX 14-DEC-2000; 2000WO-EP012698.

XX

XX 16-DEC-1999; 99US-0171121P.

XX

XX (BADI) BASF PLANT SCI GMBH.

XX

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;

XX PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;

XX

XX WPI; 2001-398121/42.

DR N-PSDB; AAH44231.

XX Tocopherol and carotenoid metabolism related protein (TCMRP), used to

PT produce fine chemicals, is isolated from mosses, algae, microorganisms,

PT fungi, plants, or their fragments.

XX Claim 28; Page 115-116; 133pp; English.

XX The present invention describes isolated tocopherol and carotenoid

CC metabolism related proteins (TCMRP) (I) from mosses or algae,

CC microorganisms or fungi, plants, or its fragments. (I) can be used as

CC enzymes in the production of fine chemicals or in the metabolism of

CC tocopherols and carotenoids. (II) also assist in transmembrane transport.

CC The fine chemicals that can be produced include lipids, fatty acids,

CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.

CC Nucleotide sequences, proteins, vectors and host cells from the present

CC invention can be used: (a) to identify mosses related to Physcomitrella

CC patens; (b) in mapping genomes of mosses related to Physcomitrella patens

CC; (c) in the modulation of TCMRP activity; (d) in evolutionary studies;

CC (e) in the determination of functional TCMRP regions; (f) and in the

CC cellular production of fine chemicals. AAH44222 to AAH44262 encode the

CC Physcomitrella patens TCMRP proteins given in AAB99849 to AAB99889.

CC AAH44212 to AAH44221 represent nucleotide sequence used in the

CC exemplification of the present invention

XX Sequence 153 AA;

SQ Query Match 91.7%; Score 22; DB 4; Length 153;

Best Local Similarity 80.0%; Pred. No. 9e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

DB 140 GYTVE 144

RESULT 30

ABP26146

ID ABP26146 standard; protein; 153 AA.

XX AC ABP26146;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 1468.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX group A streptococcus; Streptococcus pyogenes; antibacterial;

XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus pyogenes.

XX WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB004789.

XX PR 27-OCT-2000; 2000GB-00026333.

XX PR 24-NOV-2000; 2000GB-00028727.

XX PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Masighani V, Margarit Y RosI, Grandi G, Fraser C;

XX Tettelin H;

XX WPI; 2002-352536/38.

XX N-PSDB; ABN66777.

XX New Streptococcus protein for the treatment or prevention of infection or

XX disease caused by Streptococcus bacteria, such as meningitis, and for

PT

PT detecting a compound that binds to the protein.

XX Claim 1; Page 3301; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in

CC the specification. The proteins have antibacterial and antiinflammatory

CC activity. (I) nucleic acids encoding (I), ABN6044-ABN71526 and

CC antibodies that bind (I) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to

CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins

XX Sequence 153 AA;

SQ Query Match 91.7%; Score 22; DB 5; Length 153;

Best Local Similarity 80.0%; Pred. No. 9e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

DB 93 GYTVE 97

RESULT 31

ABO76341

ID ABO76341 standard; protein; 161 AA.

XX AC ABO76341;

XX DT 29-JUL-2004 (first entry)

XX DE Pseudomonas aeruginosa polypeptide #8516.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX PD 22-APR-2003.

XX PF 18-FEB-1999; 99US-00252991.

XX PR 18-FEB-1998; 98US-0074788P.

XX PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD09912.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

XX useful as molecular targets for diagnostics, prophylaxis and treatment of

XX pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 25087; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the

XX polynucleotides encoding them. The sequences are useful in diagnosis and

XX therapy of pathological conditions, as molecular targets for diagnostics,

XX prophylaxis and treatment of pathological conditions resulting from a

CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-P. aeruginosa drugs, as templates for recombinant  
 CC production of P. aeruginosa-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of P. aeruginosa-caused  
 CC infection, and in detection of P. aeruginosa sequences or other sequences  
 CC of Pseudomonas species using biochip technology. Sequences AB067826-  
 CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 161 AA;

Query Match 91.7%; Score 22; DB 7; Length 161;  
 Best Local Similarity 80.0%; Pred. No. 9.6e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVE 5  
 ||||  
 Db 54 GYSVE 58

# RESULT 32

ADH86858  
 ID ADH86858 standard; protein; 164 AA.

AC ADH86858;

DT 22-APR-2004 (first entry)

DE Enterococcus faecalis polypeptide #1338.

KW Enterococcus faecalis infection; transcription regulatory element;  
 KW antibacterial.

OS Enterococcus faecalis.

PN US6617156-B1.

PD 09-SEP-2003.

PF 13-AUG-1998; 98US-00134000.

PR 15-AUG-1997; 97US-0055778P.

PA (DOUC/) DOUCETTE-STAMM L A.

PA (BUSH/) BUSH D.

PI Doucette-Stamm LA, Bush D;

XX WPI; 2003-895394/82.

DR N-PSDB; ADH83453.

XX New nucleic acid comprising a sequence encoding an Enterococcus faecalis  
 PT polypeptide, useful for preparing a composition for diagnosing or  
 PT treating E. faecalis infection.

PS Disclosure; SEQ ID NO 4743; 193pp; English.

XX The invention relates to Enterococcus faecalis polynucleotides and  
 CC polypeptides. The invention also relates to a recombinant expression  
 CC vector comprising a polynucleotide operably linked to a transcription  
 CC regulatory element, a cell comprising a recombinant vector, a method for  
 CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising  
 CC a sequence not given in the specification, a recombinant vector  
 CC comprising the nucleic acid and a cell comprising the recombinant vector.  
 CC The polynucleotides can be used to detect the presence of E. faecalis in  
 CC a sample. The sequences are useful for preparing a composition for  
 CC diagnosing or treating Enterococcus faecalis infection. This sequence  
 CC represents an E. faecalis polypeptide of the invention.

XX

SQ Sequence 164 AA;

Query Match 91.7%; Score 22; DB 7; Length 164;  
 Best Local Similarity 80.0%; Pred. No. 9.7e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVE 5  
 ||||  
 Db 64 GYAVE 68

# RESULT 33

AAU55493  
 ID AAU55493 standard; protein; 173 AA.

XX AC AAU55493;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #16389.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US012865.

PR 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-020841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59569.

XX Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris.

PS Example 1; SEQ ID NO 16688; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX



CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 179 AA;

Query Match 91.7%; Score 22; DB 6; Length 179;  
Best Local Similarity 80.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 155 GYSVE 159

#### RESULT 36

AAG30134  
ID AAG30134 standard; protein; 180 AA.

XX AAG30134;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 35972.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; Genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 23-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

XX 05-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 06-MAY-1999; 99US-0132487P.

XX 07-MAY-1999; 99US-0132863P.

XX 11-MAY-1999; 99US-0134256P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 18-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.

```
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147132P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148694P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158233P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159333P.
PR 14-OCT-1999; 99US-0159633P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.

PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.7%; Score 22; DB 3; Length 180;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 63 GYSVE 67

RESULT 37
AAW02587
ID AAW02587 standard; protein; 184 AA.
XX
AC AAW02587;
XX
DT 28-NOV-1996 (first entry)
XX
DE Lerk-6 protein.
XX
KW Lerk-6; hek; elk; cell surface receptor; culture; reagent; neuron;
KW disorder; injury; delivery agent; diagnostic; therapeutic; Lerk-7; probe;
KW cytokine.
XX
OS Mus musculus.
XX
PN WO9617925-A1.
XX
PD 13-JUN-1996.
XX
PF 05-DEC-1995; 95WO-US015781.
XX
PR 06-DEC-1994; 94US-00351025.
PR 01-MAR-1995; 95US-00396946.
XX
PA (IMMV ) IMMUNEX CORP.
XX
PI Cerretti DP;
XX
DR WPI; 1996-287171/29.
DR N-PSDB; AAT32700.
XX
PT New isolated human Lerk-7 cytokine - which binds to cell surface
PT receptors elk, hek and eck, useful for delivering agents to cells or for
PT treating neural disorders.
XX
PS Example 1; Page 35; 49pp; English.
XX
CC The present sequence is that of a murine cytokine, Lerk-6, encoded by
CC AAT32700. A fragment of murine Lerk-6 DNA was isolated by PCR for use as
CC a probe for a human Lerk-6 homologue. The probe however led to the
CC identification of a human Lerk-7 gene (see AAT32699). The Lerk-7 gene and
CC protein can be used for studying the role of Lerk-7 in conjunction with
CC elk, hek and eck receptors. they can also be used for delivering
CC diagnostic or therapeutic agents to cells, e.g. cancer cells. The Lerk-7
CC proteins can also exhibit neuroprotective or neurotrophic properties and
CC can be used to treat neural tissue disorders
XX
SQ Sequence 184 AA;

Query Match 91.7%; Score 22; DB 2; Length 184;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
```

```

Db          29 GYVE 33
          ||||
RESULT 38
AAR91283
ID AAR91283 standard; protein; 184 AA.
XX
AC AAR91283;
XX
DT 10-OCT-1996 (first entry)
XX
DE LERK-6 protein.
XX
KW LERK-6; hek; elk; cell surface receptors; culture; reagent; neurons;
KW disorder; injury; delivery agent; diagnostic; therapeutic.
XX
OS Mus musculus.
XX
FN WO9610911-A1.
XX
PD 18-APR-1996.
XX
PF 04-OCT-1995; 95WO-US012779.
XX
PR 05-OCT-1994; 94US-00318393.
PR 03-OCT-1995; 95US-00538709.
XX
PA (IMMV ) IMMUNEX CORP.
XX
FI Cerretti DP;
XX
DR WPI; 1996-209575/21.
DR N-PSDB; AAT14009.
XX
PT Isolated DNA encoding cytokine designated LERK-6 which binds to hek and
PT elk cell surface receptors - useful for drug delivery and screening
PT procedures.
XX
PS Claim 6; Page 34; 44pp; English.
XX
CC The LERK-6 polypeptide encoded can be used to isolate cells expressing
CC hek/elk cell surface receptors, or to measure the biological activity of
CC such receptors. The protein may also be used as a delivery agent, taking
CC diagnostic and therapeutic agents to cells expressing such receptors.
CC LERK-6 can also be used as a tissue culture reagent to enhance the
CC viability or prolong the lifespan of the neurons. Neural tissue disorders
CC and injuries may be treated by contact with the polypeptide
XX
SQ Sequence 184 AA;
XX
Query Match          91.7%; Score 22; DB 2; Length 184;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 GYXVE 5
          ||||
Db          29 GYVE 33

RESULT 39
AAY06820
ID AAY06820 standard; protein; 184 AA.
XX
AC AAY06820;
XX
DT 24-JUN-1999 (first entry)
XX
DE Murine LERK-6 polypeptide.
XX
KW LERK-6 polypeptide; hek receptor; elk receptor; human; murine; injury;
KW cell proliferation; neural growth; neural tissue; neurological disease;
KW neurodegenerative; excitotoxicity.

```

```

XX OS Mus sp.
XX
XX FN WO9910495-A1.
XX
XX PD 04-MAR-1999.
XX
XX PF 27-AUG-1998; 98WO-US017772.
XX
XX PR 29-AUG-1997; 97US-00920440.
XX
XX PA (IMMV ) IMMUNEX CORP.
XX
XX FI Cerretti DP;
XX
XX WPI; 1999-243567/20.
XX N-PSDB; AAX32761.
DR
XX
XX PT New cytokine designated LERK-6.
XX
XX PS Claim 6; Page 37-38; 46pp; English.
XX
CC The invention relates murine and human LERK-6 polypeptides that bind to
CC hek/elk receptors. Host cells transfected or transformed with vectors
CC comprising the LERK-6 nucleic acid sequences are used for the recombinant
CC production of the proteins. LERK-6 polypeptides may be useful in the
CC enhancement, stimulation, proliferation or growth of cells expressing the
CC hek or elk receptor. The ligand and receptor complex may be involved in
CC neural growth, development and/or maintenance. LERK-6 can be used for
CC treating disorders of neural tissue such as injury or neurological
CC diseases, either chronic or acute. LERK-6 may be employed in treating
CC neurodegenerative conditions where there is neural death, excitotoxicity.
CC In addition, they may be administered to a mammal to exert a trophic
CC effect on neural tissue. They can also be used as reagents for those
CC conducting quality assurance studies e.g. to monitor shelf life and
CC stability of elk protein under different conditions. The polypeptides can
CC also be used as carriers for delivering agents attached to cells bearing
CC the elk or hek cell surface receptor. The present sequence represents a
CC murine LERK-6 polypeptide
XX
XX SQ Sequence 184 AA;
XX
Query Match          91.7%; Score 22; DB 2; Length 184;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 GYXVE 5
          ||||
Db          29 GYVE 33

RESULT 40
ADO06904
ID ADO06904 standard; peptide; 199 AA.
XX
AC ADO06904;
XX
XX DT 01-JUL-2004 (first entry)
XX
XX DE Lactate dehydrogenase-elevating virus gp5 protein.
XX
XX KW virucide; gene therapy; vaccine; neutralising epitope; mimotopes;
XX KW porcine reproductive and respiratory syndrome virus;
XX KW Lactate dehydrogenase-elevating virus; LDV; gp5.
XX
XX OS Lactate dehydrogenase-elevating virus.
XX
XX FN US2004014028-A1.
XX
XX PD 22-JAN-2004.
XX
XX PF 19-JUL-2002; 2002US-00199545.
XX
XX

```

PR 19-JUL-2002; 2002US-00199545.  
 XX (WOLF-) WOLF BIOTECH.  
 XX Lopez OJ, Ostrowski M;  
 XX WPI; 2004-121549/12.  
 DR Isolating neutralizing epitopes of a pathogen for vaccination purposes  
 XX comprises obtaining sera with and without neutralizing antibodies from a  
 PT species and using the difference between the sera to isolate neutralizing  
 PT epitopes.  
 XX Example 8; Fig 10; 28pp; English.  
 XX The invention describes isolating neutralising epitopes of a pathogen  
 CC comprising obtaining sera with and without neutralising antibodies from a  
 CC species, and using the difference between the sera with neutralising  
 CC antibodies and the sera without neutralising antibodies to isolate  
 CC neutralising epitopes and/or mimotopes. The composition and methods are  
 CC useful in diagnosing, preventing or treating Porcine reproductive and  
 CC respiratory syndrome virus (PRRSV) infection. The polyclonal or  
 CC monoclonal antibodies specific to epitope B, or the monospecific chimeric  
 CC mouse/porcine antibodies obtained by recombinant DNA technology specific  
 CC to epitope B, are used to facilitate cure of infected animals or  
 CC prevention of animals at risk of infection with PRRSV. This is the amino  
 CC acid sequence of lactate dehydrogenase-elevating virus (LDV) gp5 used in  
 CC a comparison with PRRSV gp5 to determine the neutralising epitope  
 CC sequences of gp5.  
 XX Sequence 199 AA;  
 SQ Query Match 91.7%; Score 22; DB 8; Length 199;  
 Best Local Similarity 80.0%; Pred. NO. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 61 GYAVE 65  
 RESULT 41  
 ABG01444  
 ID ABG01444 standard; protein; 200 AA.  
 XX AC ABG01444;  
 XX 13-FEB-2002 (first entry)  
 XX Novel human diagnostic protein #1435.  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX Homo sapiens.  
 XX WO200175067-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US008631.  
 XX 31-MAR-2000; 2000US-00540217.  
 XX 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 XX N-PSDB; AAS65631.  
 XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 20; SEQ ID NO 31803; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 200 AA;  
 SQ Query Match 91.7%; Score 22; DB 4; Length 200;  
 Best Local Similarity 80.0%; Pred. NO. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 130 GYIVE 134  
 RESULT 42  
 AAB74333  
 ID AAB74333 standard; protein; 201 AA.  
 XX AC AAB74333;  
 XX 29-JUN-2001 (first entry)  
 XX Human papillomavirus 16 E2NT module.  
 DE E2NT; antiviral; warts; proliferative skin lesion; cervical cancer; HPV.  
 KW Human papillomavirus.  
 OS WO200121645-A2.  
 XX 29-MAR-2001.  
 XX 18-SEP-2000; 2000WO-GB003568.  
 XX 17-SEP-1999; 99GB-00021938.  
 XX (UYVO-) UNIV YORK.  
 XX Antson A, Maitland N;  
 XX WPI; 2001-328091/34.  
 XX Crystallized molecular complex of the nuclear phosphoprotein E2 N-  
 PT terminal module useful in drug design for the treatment of conditions  
 PT associated with human papilloma virus infection.  
 XX Disclosure; Fig 1; 87pp; English.



CC The present invention relates to a crystallized molecular complex of an  
CC E2 N-terminal module (E2NT) dimer protein or its homologue having  
CC residues vital for transcription and replication activities of the  
CC protein. The invention is used to identify antiviral agents, and to  
CC prepare treatments for conditions that arise from herpes papillomavirus  
CC infection, particularly warts, proliferative skin lesions and/or cervical  
CC cancer. The present sequence is the E2NT module of HPV16  
XX  
SQ Sequence 201 AA;

Query Match 91.7%; Score 22; DB 4; Length 201;  
Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVE 5  
|||  
Db 114 GYTV 118

RESULT 43  
AAG93162  
ID AAG93162 standard; protein; 201 AA.

AC AAG93162;

XX 26-SEP-2001 (first entry)

DT C glutamicum protein fragment SEQ ID NO: 6916.

DE Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
XX organic acid synthesis.

XX Corynebacterium glutamicum.

XX EPI108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 95JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW ) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

XX N-PSDB; AAH68381.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analyzing  
PT expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 6916; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein  
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of coryneform bacterium, measuring expression amount and analysing  
CC the expression profile or expression pattern of a gene derived from  
CC Coryneform bacterium, and identifying a homologue of a gene derived from  
CC coryneform bacterium. Coryneform bacteria are useful for producing amino  
CC acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a protein described in the  
CC exemplification of the invention. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from the European Patent Office

XX Sequence 201 AA;

Query Match 91.7%; Score 22; DB 4; Length 201;  
Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVE 5  
|||  
Db 152 GYTV 156

RESULT 44

ABO69448

ID ABO69448 standard; protein; 203 AA.

XX ABO69448;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #1623.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD03019.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 18194; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html

XX Sequence 203 AA;

Query Match 91.7%; Score 22; DB 7; Length 203;  
Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVE 5  
|||  
Db 171 GYSVE 175

RESULT 45  
 AB071975  
 ID AB071975 standard; protein; 205 AA.  
 XX  
 AC AB071975;  
 XX  
 DT 29-JUL-2004 (first entry)  
 XX  
 XX Pseudomonas aeruginosa polypeptide #4150.  
 DE  
 XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
 KW  
 XX Pseudomonas aeruginosa.  
 OS  
 XX US6551795-B1.  
 PN  
 XX 22-APR-2003.  
 PD  
 XX 18-FEB-1999; 99US-00252991.  
 PF  
 XX 18-FEB-1998; 98US-0074788P.  
 PR  
 XX 27-JUL-1998; 98US-0094190P.  
 PR  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 FA  
 XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
 PI  
 XX WPI; 2003-615309/58.  
 DR  
 XX N-PSDB; ABD05546.  
 DD  
 XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
 PT pathological conditions resulting from bacterial infection.  
 PT  
 XX Disclosure; SEQ ID NO 20721; 455pp; English.  
 PS  
 XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-P. aeruginosa drugs, as templates for recombinant  
 CC production of P. aeruginosa-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of P. aeruginosa-caused  
 CC infection, and in detection of P. aeruginosa sequences or other sequences  
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html  
 XX  
 SQ Sequence 205 AA;  
 Query Match 91.7%; Score 22; DB 7; Length 205;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 Db 123 GYTV 127  
 RESULT 46  
 AAB79546  
 ID AAB79546 standard; protein; 206 AA.  
 XX  
 AC AAB79546;  
 XX  
 XX 30-APR-2001 (first entry)  
 DT  
 XX Corynebacterium glutamicum SMP protein sequence SEQ ID NO:608.  
 DE

XX  
 KW Corynebacterium glutamicum; carbon metabolism and energy production;  
 KW SMP protein; sugar metabolism and oxidative phosphorylation protein;  
 KW fine chemical production; organic acid; proteinogenic amino acid;  
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;  
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;  
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;  
 KW diagnosis; Corynebacterium diphtheriae; evolutionary study.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 XX WO200100844-A2.  
 PN  
 XX 04-JAN-2001.  
 PD  
 XX 23-JUN-2000; 2000WO-IB000943.  
 PF  
 XX 25-JUN-1999; 99US-0141031P.  
 PR  
 XX 08-JUL-1999; 99DE-01031412.  
 PR  
 XX 08-JUL-1999; 99DE-01031413.  
 PR  
 XX 08-JUL-1999; 99DE-01031419.  
 PR  
 XX 08-JUL-1999; 99DE-01031420.  
 PR  
 XX 08-JUL-1999; 99DE-01031424.  
 PR  
 XX 08-JUL-1999; 99DE-01031428.  
 PR  
 XX 08-JUL-1999; 99DE-01031431.  
 PR  
 XX 08-JUL-1999; 99DE-01031433.  
 PR  
 XX 08-JUL-1999; 99DE-01031434.  
 PR  
 XX 08-JUL-1999; 99DE-01031510.  
 PR  
 XX 08-JUL-1999; 99DE-01031562.  
 PR  
 XX 08-JUL-1999; 99DE-01031634.  
 PR  
 XX 09-JUL-1999; 99DE-01032180.  
 PR  
 XX 09-JUL-1999; 99DE-01032227.  
 PR  
 XX 09-JUL-1999; 99DE-01032230.  
 PR  
 XX 09-JUL-1999; 99US-0143208P.  
 PR  
 XX 14-JUL-1999; 99DE-01032924.  
 PR  
 XX 14-JUL-1999; 99DE-01032973.  
 PR  
 XX 14-JUL-1999; 99DE-01033005.  
 PR  
 XX 27-AUG-1999; 99DE-01040765.  
 PR  
 XX 31-AUG-1999; 99US-0151572P.  
 PR  
 XX 03-SEP-1999; 99DE-01042076.  
 PR  
 XX 03-SEP-1999; 99DE-01042079.  
 PR  
 XX 03-SEP-1999; 99DE-01042086.  
 PR  
 XX 03-SEP-1999; 99DE-01042087.  
 PR  
 XX 03-SEP-1999; 99DE-01042088.  
 PR  
 XX 03-SEP-1999; 99DE-01042095.  
 PR  
 XX 03-SEP-1999; 99DE-01042123.  
 PR  
 XX 03-SEP-1999; 99DE-01042125.  
 XX  
 (BADI ) BASF AG.  
 FA  
 XX Pompejus M, Kroeger B, Schroeder H, Zeider O, Haberhauer G;  
 XX  
 PI WPI; 2001-061975/07.  
 XX  
 DR N-PSDB; AAF71663.  
 DR  
 XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar  
 PT metabolism and oxidative phosphorylation protein for production or  
 PT modulation of production of fine chemicals e.g. amino acids,  
 PT carbohydrates or enzymes.  
 PT  
 XX Claim 20; Page 1014-1015; 1246pp; English.  
 PS  
 XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar  
 CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243  
 CC to AAB 79633 which are involved in carbon metabolism and energy  
 CC production. The C. glutamicum SMP gene can be used in vectors (II) for  
 CC expression in host cells and production or modulation of production of  
 CC fine chemicals, such as, an organic acid, a proteinogenic or  
 CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a  
 CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,  
 CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a  
 CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)  
 CC encoded by them are used for diagnosing the presence or activity of

CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells  
 CC containing them are used to map genomes of organisms related to C.  
 CC glutamicum, identify and localise C. glutamicum sequences of interest, in  
 CC evolutionary studies, in determining SMP protein regions required for  
 CC function, in modulating SMP protein activity, in modulating the  
 CC metabolism of sugars, and in modulating high-energy molecule production  
 CC in a cell (i.e. ATP, NADPH)  
 XX  
 SQ Sequence 206 AA;  
 Query Match 91.7%; Score 22; DB 4; Length 206;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 GYXVE 5  
 Db 37 GYAVE 41  
 RESULT 47  
 AAR94766  
 ID AAR94766 standard; protein; 209 AA.  
 XX  
 AC AAR94766;  
 XX  
 DT 02-JUL-1996 (first entry)  
 XX  
 DE Mouse EPH receptor ligand Elf-1.  
 XX  
 KW Elf-1; EPH receptor ligand; dementia; tachycardia; therapy; diagnosis;  
 KW transgenic animal.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..20  
 FT /label= Sig\_peptide  
 FT Protein 21..209  
 FT /label= Mat\_protein  
 FT Region 35..166  
 FT /label= Core\_sequence\_motif  
 FT Modified-site 38  
 FT /note= "potential N-glycosylation site"  
 FT Region 69..159  
 FT /label= Cys4\_motif  
 FT Modified-site 170  
 FT /note= "potential N-glycosylation site"  
 FT Modified-site 184  
 FT /note= "potential N-glycosylation site"  
 XX  
 PN WO9609384-A1.  
 XX  
 XX 28-MAR-1996.  
 XX  
 PF 19-SEP-1995; 95WO-US011869.  
 XX  
 PR 19-SEP-1994; 94US-00308814.  
 PR 27-FEB-1995; 95US-00393462.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Flanagan JG, Cheng H;  
 DR WPI; 1996-188446/19.  
 DR N-PSDB; AAT15008.  
 XX  
 PT Murine and chicken EPH receptor ligand, Elf-1 - useful in diagnosis and  
 PT treatment of disorders associated with the Elf-1 gene, e.g. dementia,  
 PT tachycardia , etc.  
 XX  
 PS Claim 1; Page 86; 107pp; English.  
 XX  
 CC A novel mouse EPH receptor ligand, Elf-1 (AAR94766), is involved in the

CC formation and maintenance of ordered spatial arrangements of  
 CC differentiated tissue. It is the product of a cDNA clone (AAT15008) obt'd.  
 CC from an embryo mid- and hind-brain cDNA expression library. This cDNA is  
 CC used for the prodn. of recombinant Elf-1, which can be used to modulate  
 CC proliferation, survival and/or differentiation of cells and tissues, and  
 CC to stimulate or antagonise intracellular signal transduction pathways  
 CC mediated by the EPH-type receptor  
 XX  
 SQ Sequence 209 AA;  
 Query Match 91.7%; Score 22; DB 2; Length 209;  
 Best Local Similarity 80.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 GYXVE 5  
 Db 54 GYTVE 58  
 RESULT 48  
 AAW71006  
 ID AAW71006 standard; protein; 209 AA.  
 XX  
 AC AAW71006;  
 XX  
 DT 20-OCT-1998 (first entry)  
 XX  
 DE Amino acid sequence of a mammalian Elf-1 protein.  
 XX  
 KW Mouse; EPH receptor ligand; Elf-1; mek-4; sek-AP; tyrosine kinase ligand;  
 KW B61; LERK-2; proliferation; differentiation; intracellular signalling;  
 KW increased; survival; neuronal cell; neuron survival; treatment;  
 KW Alzheimer's; Parkinson's; lymphatic tumour; artificial liver; cartilage;  
 KW bone formation.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..20  
 FT /note= "signal peptide"  
 FT Protein 21..209  
 FT /note= "mature protein"  
 FT Misc-difference 35..166  
 FT /note= "contains core sequence motif"  
 FT Modified-site 38  
 FT /note= "potential N-linked glycosylation site"  
 FT Misc-difference 69..159  
 FT /note= "contains a Cys4 motif"  
 FT Modified-site 170  
 FT /note= "potential N-linked glycosylation site"  
 FT Modified-site 184  
 FT /note= "potential N-linked glycosylation site"  
 XX  
 PN US5795734-A.  
 XX  
 XX 18-AUG-1998.  
 XX  
 PF 31-MAY-1995; 95US-00455001.  
 XX  
 PR 19-SEP-1994; 94US-00308814.  
 PR 27-FEB-1995; 95US-00393462.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Flanagan JG, Cheng H;  
 XX  
 XX WPI; 1998-466665/40.  
 DR N-PSDB; AAV42926.  
 XX  
 PT Nucleic acid encoding Elf-1 protein that binds to EPH-type receptor - for  
 PT production of Elf-1 protein, useful for regulating proliferation,  
 PT differentiation, and survival of cells.  
 FT  
 XX

PS Claim 1; Fig 2A; 53pp; English.

XX The present sequence represents a mammalian EPH receptor ligand designated E1f-1. This ligand can bind to both mex-4 and sek-AP. E1f-1 is a tyrosine kinase ligand, which is linked to the membrane through a phosphatidylinositol linkage. It shares some homology to 2 other EPH receptor ligands, B61 and LERK-2. The E1f-1 protein modulates proliferation, differentiation and survival of EPH receptor-expressing cells by stimulating or antagonising intracellular signalling mediated by the EPH receptor. Typical of many potential applications are increasing survival of neuronal cells in culture (e.g. where intended for transplantation), also therapeutically in increase neuron survival (e.g. treatment of Alzheimer's or Parkinson's diseases), to prevent nervous system and lymphatic tumours, to induce differentiation of hepatocytes to form an artificial liver, to induce cartilage and bone formation

XX SQ Sequence 209 AA;

Query Match 91.7%; Score 22; DB 2; Length 209;  
Best Local Similarity 80.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GYXVE 5  
|||

DB 54 GYTVE 58

RESULT 49  
AA06822  
ID AAY06822 standard; protein; 213 AA.

XX AC AAY06822;

XX DT 24-JUN-1999 (first entry)

XX DE Human LERK-6 polypeptide.

XX LERK-6 polypeptide; hek receptor; elk receptor; human; murine; injury;  
KW cell proliferation; neural growth; neural tissue; neurological disease;  
KW neurodegenerative; excitotoxicity.

XX OS Homo sapiens.

XX PN WO9910495-A1.

XX PD 04-MAR-1999.

XX PF 27-AUG-1998; 98WO-US017772.

XX PR 29-AUG-1997; 97US-00920440.

XX PA (IMMV ) IMMUNEX CORP.

XX PI Cerretti DP;

XX WPI: 1999-243567/20.  
DR N-PSDB; AAX32767.

XX PT New cytokine designated LERK-6.

XX PS Claim 6; Page 42; 46pp; English.

XX The invention relates murine and human LERK-6 polypeptides that bind to hek/elk receptors. Host cells transfected or transformed with vectors comprising the LERK-6 nucleic acid sequences are used for the recombinant production of the proteins. LERK-6 polypeptides may be useful in the enhancement, stimulation, proliferation or growth of cells expressing the hek or elk receptor. The ligand and receptor complex may be involved in neural growth, development and/or maintenance. LERK-6 can be used for treating disorders of neural tissue such as injury or neurological diseases, either chronic or acute. LERK-6 may be employed in treating neurodegenerative conditions where there is neural death, excitotoxicity. In addition, they may be administered to a mammal to exert a trophic

CC effect on neural tissue. They can also be used as reagents for those conducting quality assurance studies e.g. to monitor shelf life and stability of elk protein under different conditions. The polypeptides can also be used as carriers for delivering agents attached to cells bearing the elk or hek cell surface receptor. The present sequence represents a human LERK-6 polypeptide

XX SQ Sequence 213 AA;

Query Match 91.7%; Score 22; DB 2; Length 213;  
Best Local Similarity 80.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GYXVE 5  
|||

DB 58 GYTVE 62

RESULT 50  
AA014749  
ID AAO14749 standard; protein; 224 AA.

XX AC AAO14749;

XX DT 28-JUN-2002 (first entry)

XX DE Corynebacterium glutamicum sigma factor M protein.

XX SigM gene; sigma factor M; hybridisation probe; L-amino acid preparation;  
KW L-lysine.

XX OS Corynebacterium glutamicum.

XX PN WO200218599-A1.

XX PD 07-MAR-2002.

XX PF 30-AUG-2001; 2001WO-EP009972.

XX PR 02-SEP-2000; 2000DE-01043337.  
PR 28-JUL-2001; 2001DE-01036984.

XX FA (DEGS ) DEGUSSA AG.

XX Bathe B, Bastuck C, Farwick M, Hermann T, Pfefferle W;  
XX WPI: 2002-315544/35.  
DR N-PSDB; AAL42355.

XX New sigM gene from coryneform bacteria useful as probe to isolate genes which code for sigma factor M, and overexpression of which gene in coryneform bacteria is useful for producing amino acids, especially L-lysine.

XX Claim 7; Page 38-39; 42pp; English.

XX The invention comprises a Corynebacterium glutamicum sigma factor M (sigM) protein and coding sequence. The sigM gene sequence is useful as a hybridisation probe for discovering RNA, cDNA and DNA that code for sigM, or have a high similarity with sigM sequences. Corynebacterium in which the sigM gene is enhanced, is useful for preparing L-amino acids (i.e. L-lysine). The present amino acid sequence represents the Corynebacterium glutamicum sigM protein

XX SQ Sequence 224 AA;

Query Match 91.7%; Score 22; DB 5; Length 224;  
Best Local Similarity 80.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GYXVE 5  
|||

DB 175 GYTVE 179

```

RESULT 51
ID ADC96464 standard; protein; 232 AA.
XX
AC ADC96464;
XX
XX 01-JAN-2004 (first entry)
XX
DE E. faecium protein sequence SEQ ID 6091.
XX
KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.
XX
OS Enterococcus faecium.
XX
PN US6583275-B1.
XX
PD 24-JUN-2003.
XX
PF 30-JUN-1998; 98US-00107532.
XX
PR 02-JUL-1997; 97US-0051571P.
PR 14-MAY-1998; 98US-0085598P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
XX WPI; 2003-799836/75.
DR N-PSDB; ADC92810.
XX
PT New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.
XX
PS Example 1; SEQ ID NO 6091; 243pp; English.
XX
CC The invention relates to an isolated nucleic acid derived from
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridising to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC transcription regulatory element, a cell comprising the vector and a
CC single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids are useful for diagnosing pathological conditions
CC resulting from E. faecium bacterial infection (e.g. urinary tract
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of Candida albicans -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating Enterococcus faecium infections. The present sequence represents
CC one if the disclosed E. faecium proteins.
XX
SQ Sequence 232 AA;

Query Match 91.7%; Score 22; DB 7; Length 232;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 67 GYTV 71

RESULT 52
ID ABM68208 standard; protein; 235 AA.
XX

```

```

XX ABM68208;
AC
XX 20-NOV-2003 (first entry)
XX
DE Photorhabdus luminescens protein sequence #1305.
XX
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough.
XX
OS Photorhabdus luminescens.
XX
PN WO200294867-A2.
XX
PD 28-NOV-2002.
XX
PF 07-FEB-2002; 2002WO-IB003040.
XX
PR 07-FEB-2001; 2001FR-00001659.
XX
PA (INSP ) INST PASTEUR.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI Buchrieser C;
XX
XX WPI; 2003-148459/14.
XX
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
XX Claim 2; SEQ ID NO 1305; 1205pp; French.
XX
CC The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens proteins
XX
SQ Sequence 235 AA;

Query Match 91.7%; Score 22; DB 6; Length 235;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 60 GYTV 64

RESULT 53
ID ABG91554 standard; protein; 238 AA.
XX
AC ABG91554;

```

XX 18-NOV-2002 (first entry)  
 XX Purine/pyrimidine triphosphate type nucleotidyltransferase #139.  
 XX Nucleotidyltransferase; enzyme; active site engineering;  
 KW alpha-D-glucopyranosyl phosphate thymidyltransferase; Ep;  
 KW substrate specificity; nucleotide sugar;  
 KW glycosylated bioactive natural product.  
 XX Halobacterium sp. NRC-1.  
 XX WO200248331-A2.  
 XX 20-JUN-2002.  
 XX 13-DEC-2001; 2001WO-US047953.  
 XX 13-DEC-2000; 2000US-0254927P.  
 XX (SLOK ) SLOAN KETTERING INST CANCER RES.  
 XX Thorson JS, Nikilov DB;  
 XX WPI; 2002-608282/65.  
 XX Nucleotidyltransferase mutated at one or more amino acids, useful in  
 PT the synthesis of nucleotide sugars.  
 XX Claim 3; Page; 182pp; English.  
 XX The invention relates to a Nucleotidyltransferase mutated at one or  
 CC more amino acids selected from V173, G147, W224, N112, G175, D111, E162,  
 CC T201, T200, E199, L195, L89, L89T, L109, Y146 or Y177 (with reference to  
 CC the Salmonella enterica rnaA-encoded alpha-D-glucopyranosyl phosphate  
 CC thymidyltransferase, Ep, enzyme appearing as ABG91798). The mutations  
 CC alter the substrate specificity of the enzymes. The mutants and methods  
 CC involving them are used in the synthesis of nucleotide sugars for  
 CC altering nucleotidyltransferase substrate specificity. The  
 CC nucleotidyltransferase exhibits different substrate specificity for  
 CC GTP, CTP, TTP, UTP and ATP than a non-mutated nucleotidyltransferase.  
 CC The mutant may also exhibit a high degree of sequence identity to  
 CC Salmonella enterica LT2 alpha-D-glucopyranosyl phosphate  
 CC thymidyltransferase (Ep) and can convert a wide variety of phosphates.  
 CC The mutants can be exploited in the biosynthesis of glycosylated  
 CC bioactive natural products of pharmacological use. The present sequence  
 CC is a nucleotidyltransferase exhibiting a high degree of sequence  
 CC identity to Salmonella enterica LT2 alpha-D-glucopyranosyl phosphate  
 CC thymidyltransferase (Ep). Note: The present sequence is not displayed  
 CC in the specification but was obtained from Genbank  
 XX  
 XX Sequence 238 AA;  
 Query Match 91.7%; Score 22; DB 5; Length 238;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 206 GYSVE 210  
 RESULT 54  
 ADM25747  
 ID ADM25747 standard; protein; 238 AA.  
 AC ADM25747;  
 XX 20-MAY-2004 (first entry)  
 XX Hyperthermophile Methanopyrus kandleri protein #353.  
 DE Hyperthermophile; protein stability enhancement;  
 KW

KW protein activity enhancement.  
 XX Methanopyrus kandleri.  
 XX WO2003076575-A2.  
 XX 18-SEP-2003.  
 XX 04-MAR-2003; 2003WO-US006664.  
 XX 04-MAR-2002; 2002US-0361742P.  
 PR 14-MAY-2002; 2002US-0380423P.  
 PR 16-SEP-2002; 2002US-0410974P.  
 XX (FIDE-) FIDELITY SYSTEMS INC.  
 PA (MALY/) MALYKH A.  
 XX Slesarev AI, Pavlov A, Pavlova N, Kozyavkin S;  
 XX WPI; 2003-748383/70.  
 DR N-PSDB; ADM27081.  
 XX New isolated nucleic acids encoding any of about 1700 Methanopyrus  
 PT kandleri proteins, and the encoded proteins, useful as a medicaments or  
 PT as diagnostic agents.  
 XX Claim 31; SEQ ID NO 353; 1023pp; English.  
 XX The invention comprises the amino acid sequence of proteins from the  
 CC hyperthermophile Methanopyrus kandleri, the invention also comprises the  
 CC complete genome from Methanopyrus kandleri. The Methanopyrus kandleri  
 CC proteins of the invention are useful for enhancing the stability and/or  
 CC activity of other proteins. The Methanopyrus kandleri genome is useful in  
 CC a variety of diagnostic and analytical methods. The present amino acid  
 CC sequence represents a Methanopyrus kandleri protein of the invention.  
 XX Sequence 238 AA;  
 Query Match 91.7%; Score 22; DB 7; Length 238;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 11 GYSVE 15  
 RESULT 55  
 AAW99400  
 ID AAW99400 standard; protein; 246 AA.  
 XX AAW99400;  
 XX 08-JUN-1999 (first entry)  
 XX S. antibioticus oleM gene product.  
 DE Gene cluster; bacterium; enzyme; macrolide; antibiotic; oleandomycin;  
 KW secondary metabolite; hybridisation; probe; glycosylation; macrolactone.  
 XX Streptomyces antibioticus.  
 XX WO9905283-A2.  
 XX 04-FEB-1999.  
 XX 21-JUL-1998; 98WO-FR001593.  
 XX 25-JUL-1997; 97FR-00009458.  
 PR 12-JUN-1998; 98FR-00007411.  
 XX (HMRI ) HOECHST MARION ROUSSEL.  
 PA

PI Fromentin C, Michel J, Raynal M, Salah-Bey K, Cortes J;  
 XX Gaisser S, Leadlay P, Mendez C, Salas JA;  
 DR WPI; 1999-142938/12.  
 XX N-PSDB; AAX25776.  
 XX New nucleic acid sequences encoding enzymes involved in macrolide  
 PT biosynthesis - useful for producing hybrid secondary metabolites,  
 PT particularly erythromycin analogues.  
 XX Disclosure; Fig. 22; 221pp; French.  
 XX This sequence represents the product encoded by the oleM gene from the  
 CC olepi-olef gene cluster from the bacterium Streptomyces antibioticus. The  
 CC olef gene cluster encodes enzymes involved in the production of the  
 CC macrolide antibiotic oleandomycin as a secondary metabolite. The genes  
 CC are used to produce hybrid secondary metabolites in *S. antibioticus*, i.e.  
 CC oleandomycin analogues which may have improved properties or as  
 CC hybridisation probes for isolating homologous genes involved in  
 CC glycosylation of macrolactones in macrolide-producing strains  
 XX Sequence 246 AA;  
 SQ

Query Match 91.7%; Score 22; DB 2; Length 246;  
 Best Local Similarity 80.0%; Pred. NO. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 225 GYTV 229

RESULT 56  
 ADA34441  
 ID ADA34441 standard; protein; 250 AA.  
 XX

AC ADA34441;  
 XX 20-NOV-2003 (first entry)  
 XX Acinetobacter baumannii protein #1602.  
 DE

XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
 KW plant biocontrol agent.  
 XX

OS Acinetobacter baumannii.  
 XX

PN US6562958-B1.  
 XX

PD 13-MAY-2003.  
 XX

PF 04-JUN-1999; 99US-00328352.  
 XX

PR 09-JUN-1998; 98US-0088701P.  
 XX

PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX

PI Breton G, Bush D;  
 XX

XX WPI; 2003-576092/54.  
 DR

DR N-PSDB; ADA30315.  
 XX

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
 PT for diagnosing a bacterial disease, as components of antibacterial  
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
 PT plants.  
 XX

PS Example; SEQ ID NO 5728; 328pp; English.  
 XX

XX The invention relates to isolated Acinetobacter baumannii nucleic acids.  
 CC

CC The A. baumannii nucleic acids and polypeptides are useful as reagents  
 CC for diagnosing a bacterial disease, as components of antibacterial

CC vaccines, as targets for antibacterial drugs, to detect the presence of

CC A. baumannii and other Acinetobacter species in a sample, in screening  
 CC compounds for the ability to interfere with the A. baumannii life cycle  
 CC or to inhibit A. baumannii infection, and as biocontrol agents for  
 CC plants. The present sequence represents the amino acid sequence of an A.  
 CC baumannii protein.  
 XX

SQ Sequence 250 AA;  
 Query Match 91.7%; Score 22; DB 6; Length 250;  
 Best Local Similarity 80.0%; Pred. NO. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 136 GYSVE 140

RESULT 57

ABB61132  
 ID ABB61132 standard; protein; 264 AA.  
 XX

AC ABB61132;  
 XX

DT 26-MAR-2002 (first entry)  
 XX

DE Drosophila melanogaster polypeptide SEQ ID NO 10188.  
 XX

KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX

OS Drosophila melanogaster.  
 XX

PN WO200171042-A2.  
 XX

PD 27-SEP-2001.  
 XX

PF 23-MAR-2001; 2001WO-US009231.  
 XX

PR 23-MAR-2000; 2000US-0191637P.  
 XX

PR 11-JUL-2000; 2000US-00614150.  
 XX

PA (PEKE ) PE CORP NY.  
 XX

PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX

XX WPI; 2001-656860/75.  
 DR

DR N-PSDB; ABL05235.  
 XX

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX

PS Disclosure; SEQ ID NO 10188; 21pp + Sequence Listing; English.  
 XX

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511); expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 264 AA;  
 SQ

Query Match 91.7%; Score 22; DB 4; Length 264;  
 Best Local Similarity 80.0%; Pred. NO. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||

Db 197 GYTV 201

RESULT 58  
AAO24038  
ID AAO24038 standard; protein; 270 AA.  
XX  
AC AAO24038;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Aspergillus foetidus lipase protein.  
XX  
KW Fungal; enzyme; protein co-ordinate data; lipolytic; phospholipase; PL;  
KW lysophospholipase; LPL; dough; baked product; baking performance.  
XX  
OS Aspergillus foetidus.  
XX  
PN WO2003060112-A1.  
XX  
PD 24-JUL-2003.  
XX  
PF 16-JAN-2003; 2003WO-DK0000028.  
XX  
PR 16-JAN-2002; 2002DK-000000074.  
XX  
PA (NOVO ) NOVOZYMES AS.  
XX  
PI Svendsen A, Vind J, Heldt-Hansen HP, Christiansen L;  
XX  
DR WPI; 2003-627384/59.  
XX  
PT Producing lipolytic enzyme variants that have high activity for ester  
PT bond in amphiphilic substrate with two lipophilic groups and low activity  
PT for ester bond in amphiphilic substrate with one lipophilic group.  
XX  
PS Disclosure; Fig 2; 80pp; English.  
XX  
CC The invention relates to a novel method which comprises producing a  
CC lipolytic enzyme variant. The method of the invention may be useful for  
CC producing a lipolytic enzyme variant with a relatively high phospholipase  
CC (PL) activity and a relatively low lysophospholipase (LPL) activity. The  
CC lipolytic enzyme variant may be useful for preparing a dough or a baked  
CC product made from dough. The variant enzyme demonstrates improved baking  
CC performance, generating a product with a lower dough stickiness, a better  
CC dough extensibility and elasticity, improved dough stability and crumb  
CC structure of the baked product, a larger loaf volume and/or improved  
CC resistance to over-proofing or other abuse. The current sequence is that  
CC of the Aspergillus foetidus lipase protein of the invention  
XX  
SQ Sequence 270 AA;  
PS  
XX  
CC The invention relates to a novel method which comprises producing a  
CC lipolytic enzyme variant. The method of the invention may be useful for  
CC producing a lipolytic enzyme variant with a relatively high phospholipase  
CC (PL) activity and a relatively low lysophospholipase (LPL) activity. The  
CC lipolytic enzyme variant may be useful for preparing a dough or a baked  
CC product made from dough. The variant enzyme demonstrates improved baking  
CC performance, generating a product with a lower dough stickiness, a better  
CC dough extensibility and elasticity, improved dough stability and crumb  
CC structure of the baked product, a larger loaf volume and/or improved  
CC resistance to over-proofing or other abuse. The current sequence is that  
CC of the Aspergillus foetidus lipase protein of the invention  
XX  
SQ Sequence 270 AA;  
Query Match 91.7%; Score 22; DB 7; Length 270;  
Best Local Similarity 80.0%; Pred. No. 1.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
DB 163 GYSVE 167

RESULT 59  
AAO24039  
ID AAO24039 standard; protein; 270 AA.  
XX  
AC AAO24039;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Aspergillus niger lipase protein 2.  
XX  
KW Fungal; enzyme; protein co-ordinate data; lipolytic; phospholipase; PL;  
KW lysophospholipase; LPL; dough; baked product; baking performance.

XX  
OS Aspergillus niger.  
XX  
PN WO2003060112-A1.  
XX  
PD 24-JUL-2003.  
XX  
PF 16-JAN-2003; 2003WO-DK0000028.  
XX  
PR 16-JAN-2002; 2002DK-000000074.  
XX  
PA (NOVO ) NOVOZYMES AS.  
XX  
PI Svendsen A, Vind J, Heldt-Hansen HP, Christiansen L;  
XX  
DR WPI; 2003-627384/59.  
XX  
PT Producing lipolytic enzyme variants that have high activity for ester  
PT bond in amphiphilic substrate with two lipophilic groups and low activity  
PT for ester bond in amphiphilic substrate with one lipophilic group.  
XX  
PS Disclosure; Fig 2; 80pp; English.  
XX  
CC The invention relates to a novel method which comprises producing a  
CC lipolytic enzyme variant. The method of the invention may be useful for  
CC producing a lipolytic enzyme variant with a relatively high phospholipase  
CC (PL) activity and a relatively low lysophospholipase (LPL) activity. The  
CC lipolytic enzyme variant may be useful for preparing a dough or a baked  
CC product made from dough. The variant enzyme demonstrates improved baking  
CC performance, generating a product with a lower dough stickiness, a better  
CC dough extensibility and elasticity, improved dough stability and crumb  
CC structure of the baked product, a larger loaf volume and/or improved  
CC resistance to over-proofing or other abuse. The current sequence is that  
CC of the Aspergillus niger lipase protein 2 of the invention  
XX  
SQ Sequence 270 AA;  
Query Match 91.7%; Score 22; DB 7; Length 270;  
Best Local Similarity 80.0%; Pred. No. 1.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
DB 163 GYSVE 167

RESULT 60  
AAW18875  
ID AAW18875 standard; protein; 277 AA.  
XX  
AC AAW18875;  
XX  
DT 25-MAR-2003 (revised)  
DT 04-NOV-1997 (first entry)  
XX  
DE Babesia equi merozoite protein.  
XX  
KW Babesia equi; merozoite; conserved epitope; monoclonal antibody;  
KW Mab 36/133.97; immunoassay reagent; detection; diagnosis; infection;  
KW horse; competitive inhibition enzyme linked immunosorbent assay.  
XX  
OS Babesia equi.  
XX  
PN US5643737-A.  
XX  
PD 01-JUL-1997.  
XX  
PF 02-MAR-1995; 95US-00400413.  
XX  
PR 22-JUN-1992; 92US-00902164.  
XX  
PA (USDA ) US SEC OF AGRIC.  
XX



PI Perryman LE, Knowles DP;  
 XX WPI; 1997-350234/32.  
 XX Babesia equi merozoite protein - useful as immunoassay reagent for  
 PT detecting antibodies to B. equi in horse serum.  
 XX Claim 2; Col 19; 18pp; English.  
 XX This is a Babesia equi merozoite protein which contains a conserved  
 CC epitope specifically bound by monoclonal antibody Mab 36/133.97 (ATCC  
 CC 11788). The protein is used as an immunoassay reagent for detecting  
 CC antibodies to B. equi in serum samples from horses, especially in a  
 CC competitive immunoassay using a monoclonal antibody and a recombinant  
 CC form of the protein. (Updated on 25-MAR-2003 to correct PF field.)  
 XX SQ Sequence 277 AA;  
 Query Match 91.7%; Score 22; DB 2; Length 277;  
 Best Local Similarity 80.0%; Pred. NO. 1.7e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 64 GYAVE 68  
 RESULT 61  
 ID AAY13560 standard; protein; 282 AA.  
 XX AAY13560;  
 XX 30-JUL-1999 (first entry)  
 XX S. haemolyticus D-amino acid aminotransferase.  
 XX Listeria; vaccine; auxotrophic; attenuated; mutation; dat gene; dal gene;  
 KW T-cell response; HIV-1 antigen; HIV; aminotransferase.  
 XX Staphylococcus haemolyticus.  
 XX WO9925376-A1.  
 XX 27-MAY-1999.  
 XX 13-NOV-1998; 98WO-US024357.  
 XX 18-NOV-1997; 97US-00972902.  
 XX (UYPE-) UNIV PENNSYLVANIA.  
 XX Frankel FR, Portnoy DA;  
 XX WPI; 1999-337876/28.  
 XX Attenuated Listeria strain which is auxotrophic for D-alanine synthesis.  
 XX Disclosure; Fig 4A-B; 67pp; English.  
 XX The invention provides a vaccine comprising an auxotrophic attenuated  
 CC strain of Listeria which expresses an antigen, where the strain comprises  
 CC a mutation in at least one gene essential for growth of the Listeria  
 CC especially that the strain is auxotrophic for D-alanine. The gene is  
 CC selected from the Listeria dat and dal genes. The vaccine can be used in  
 CC mammals to stimulate T-cell responses to an antigen, especially an HIV-1  
 CC antigen. This is useful in creating effective vaccines especially against  
 CC HIV. The attenuated Listeria strain can be used as a bacterial vaccine  
 CC vector to raise an immunoresponse in e.g. immunocompromised patient or a  
 CC pregnant patient. The present sequence represents a deduced D-amino acid  
 CC aminotransferase of S. haemolyticus  
 XX SQ Sequence 282 AA;

Query Match 91.7%; Score 22; DB 2; Length 282;  
 Best Local Similarity 80.0%; Pred. NO. 1.8e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVE 5  
 DB 64 GYTVB 68  
 RESULT 62  
 ID ABP66166 standard; protein; 283 AA.  
 XX ABP66166;  
 XX 19-NOV-2002 (first entry)  
 XX Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:910.  
 XX Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;  
 KW antidiarrhetic; antibacterial; inhibitor of Salmonella; detection;  
 KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;  
 KW rotavirus; food composition; pharmaceutical composition.  
 XX Bifidobacterium longum.  
 XX EP1227152-A1.  
 XX 31-JUL-2002.  
 XX 30-JAN-2001; 2001EP-00102050.  
 XX 30-JAN-2001; 2001EP-00102050.  
 XX (NEST) SOC PROD NESTLE SA.  
 XX WPI; 2002-668397/72.  
 XX Novel polynucleotide comprising Bifidobacterium genome sequence useful as  
 PT a probe or primer for detecting and/or identifying Bifidobacterium longum  
 PT in a biological sample.  
 XX Claim 3; SEQ ID NO 910; 80pp; English.  
 XX The present invention describes a polynucleotide (I) comprising a  
 CC sequence of a Bifidobacterium genome selected from the nucleotide  
 CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at  
 CC least 90% identity or which hybridises with the sequences given in  
 CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a  
 CC fusion protein, comprising a sequence selected from 1097 sequences given  
 CC in ABP65258 to ABP66354 ligated in frame to a polynucleotide encoding a  
 CC heterologous polypeptide. (I) has antidiarrhetic and antibacterial  
 CC activities, and can be used as an inhibitor of Salmonella. (I) (which is  
 CC a probe) is useful for the detection and/or identification of  
 CC Bifidobacterium longum in a biological sample. A carrier containing the  
 CC lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM 1-2618) can be  
 CC used for preventing and/or treating diarrhoea brought about by pathogenic  
 CC bacteria and/or rotavirus. The carrier is a food composition selected  
 CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented  
 CC products, ice-creams, fermented cereal based products, milk based  
 CC powders, infant formula, pet food or a pharmaceutical composition  
 CC selected from tablets, liquid bacterial suspensions, dried oral  
 CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.  
 CC (I) is useful in DNA arrays or chips to carry out analysis of the  
 CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent  
 CC Bifidobacterium related nucleotide sequences given in the Sequence  
 CC Listing from the present invention but not mentioned further within the  
 CC specification. N.B. The sequence data for this patent is not represented  
 CC in the printed specification but is based on sequence information  
 CC supplied by the European Patent Office  
 XX SQ Sequence 283 AA;

Query Match 91.7%; Score 22; DB 5; Length 283;  
Best Local Similarity 80.0%; Pred. No. 1.8e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 75 GYAVE 79

RESULT 63  
ADF41714  
ID ADF41714 standard; protein; 296 AA.  
XX ADF41714;  
XX  
XX  
XX 12-FEB-2004 (first entry)  
XX Bacillus subtilis rocF amino acid sequence SEQ ID NO:36.  
XX  
XX Bacillus; sbo; slr; ybcO; csu; spollSA; sigB; phrC; rapA; Cess; trpA;  
KW trpB; trpC; trpD; trpE; trpF; tdn/kbl; alsD; sigD; gapB; pckA; fbp;  
KW roCA; ycgN; ycgM; rocF; rocD; enhancing expression.  
XX  
XX Bacillus subtilis.  
OS  
XX WO2003083125-A1.  
XX  
XX  
XX 09-OCT-2003.  
XX  
XX 28-MAR-2003; 2003WO-US009585.  
XX  
XX 29-MAR-2002; 2002US-0368858P.  
XX  
XX 29-MAR-2002; 2002US-0368949P.  
XX  
XX 29-APR-2002; 2002US-0376343P.  
XX  
XX (GEMV ) GENENCOR INT INC.  
XX  
XX Ferrari E, Harbison C, Rashid MH, Weyler W;  
PI  
XX  
XX WPI; 2003-876987/81.  
XX  
XX N-PSDB; ADF41713.  
XX  
XX  
XX Enhancing expression of a protein of interest from Bacillus by obtaining  
PT an altered Bacillus strain capable of producing a protein of interest and  
PT growing the altered Bacillus strain.  
XX  
XX  
XX Claim 88; SEQ ID NO 36; 114pp; English.

The present invention describes a method for enhancing the expression of a protein of interest from Bacillus. The method comprises: (a) obtaining an altered Bacillus strain capable of producing a protein of interest, where the altered Bacillus strain has at least one inactivated chromosomal gene consisting of sbo, sir, ybcO, csu, spollSA, sigB, phrC, rapA, Cess, trpA, trpB, trpC, trpD, trpE, trpF, tdn/kbl, alsD, sigD, prpC, gapB, pckA, fbp, roCA, ycgN, ycgM, rocF, and rocD; and (b) growing the altered Bacillus strain under conditions such that the protein of interest is expressed by the altered Bacillus strain, where the expression of the protein of interest is enhanced compared to the expression of the protein of interest in an unaltered Bacillus host strain. Also described: (1) an altered Bacillus strain comprising a chromosomal deletion of one or more genes consisting of sbo, sir, ybcO, csu, spollSA, sigB, phrC, rapA, Cess, trpA, trpB, trpC, trpD, trpE, trpF, tdn/kbl, alsD, sigD, prpC, gapB, pckA, fbp, roCA, ycgN, ycgM, rocF, and rocD; (2) a DNA construct comprising the gene; (3) a plasmid comprising the DNA construct; (4) a host cell comprising the plasmid; (5) a method for obtaining an altered Bacillus strain with enhanced protease production; (6) a method for enhancing expression of a protease in an altered Bacillus; (7) a method for enhancing the expression of a protein of interest in Bacillus; (8) a method for enhancing the expression of a protein of interest in Bacillus; (9) a method for obtaining a protein of interest from a Bacillus strain; (10) an isolated nucleic acid comprising a sequence having e.g., 721, 857 or 995 bp and encoding a sequence having

e.g., 211, 387 or 433 amino acids; and (11) an isolated amino acid sequence comprising e.g., 211, 387 or 433 amino acids. The method is useful for enhancing expression of a protein of interest from Bacillus. The present sequence is used in the exemplification of the present invention.

Query Match 91.7%; Score 22; DB 7; Length 296;  
Best Local Similarity 80.0%; Pred. No. 1.8e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 41 GYTV 45

RESULT 64  
AAW76698  
ID AAW76698 standard; protein; 297 AA.  
XX AAW76698;  
XX  
XX 17-OCT-2003 (revised)  
DT 28-JAN-1999 (first entry)  
XX  
XX Lipase 3 enzyme.  
XX  
XX Lipase 3 enzyme; flour dough; glycerol oxidase; bread; noodle; pasta.  
XX  
XX Aspergillus tubingensis.  
XX  
XX Key Location/Qualifiers  
FT Peptide 1..27 /note= "signal peptide"  
FT Protein 28..297 /note= "mature protein"  
FT  
XX WO9844804-A1.  
XX  
XX 15-OCT-1998.  
XX  
XX 03-APR-1998; 98WO-DK000136.  
XX  
XX 09-APR-1997; 97DK-00000400.  
XX  
XX (DANI-) DANISCO AS.  
XX  
XX Soe JB, Poulsen CH, Rasmussen P, Madrid SM, Zargahi MR;  
PI  
XX  
XX WPI; 1998-568288/48.  
XX  
XX N-PSDB; AAW62046.  
XX  
XX Improving rheological properties of flour dough - by adding glycerol oxidase and optionally lipase, used to increase dough strength, and specific volume and crumb structure of baked products.  
XX  
XX Example 6; Page 37-38; 82pp; English.  
XX  
XX The present sequence represents the Aspergillus tubingensis lipase 3 enzyme. The rheological properties of flour dough, and quality of finished products made from the dough, can be improved by adding lipase 3 together with glycerol oxidase. The doughs are particularly used to make bread, noodles and pasta. (Updated on 17-OCT-2003 to standardise OS field)  
XX  
XX Sequence 297 AA;  
SQ  
Query Match 91.7%; Score 22; DB 2; Length 297;  
Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

```
Db      190 GYSVE 194
|| ||
RESULT 65
AAW33009
ID AAW33009 standard; protein; 297 AA.
XX
AC AAW33009;
XX
DT 22-MAY-1998 (first entry)
XX
DE Aspergillus lysophospholipase.
XX
KW Lysophospholipase; maltose syrup clarity; beta-amylase;
KW maltodextrin solution.
XX
OS Aspergillus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT Peptide /label= sig_peptide
FT Peptide 28..297
FT Peptide /label= mat_peptide
XX
FN EP808903-A2.
XX
PD 26-NOV-1997.
XX
PF 24-MAR-1997; 97EP-00104976.
XX
PR 22-MAY-1996; 96DE-01020649.
XX
PA (ROHG ) ROEHM GMBH.
XX
PI Loeffler F, Khanh QN, Schuster E, Sproessler B, Wolf S, Thomas L;
XX WPI; 1998-001789/01.
DR N-PSDB; AAT88341.
XX
PT DNA encoding Aspergillus lyso:phospholipase - useful to improve
PT filtration properties of starch hydrolysates.
XX
PS Claim 16; Page 15-18; 26pp; German.
XX
CC The present sequence is an Aspergillus lysophospholipase, which can be
CC used to improve the clarity of maltose syrups produced by beta-amylase
CC treatment of maltodextrin solutions
XX
SQ Sequence 297 AA;

Query Match 91.7%; Score 22; DB 2; Length 297;
Best Local Similarity 80.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 190 GYSVE 194

RESULT 66
AAW64449
ID AAW64449 standard; protein; 297 AA.
XX
AC AAW64449;
XX
DT 16-OCT-1998 (first entry)
XX
DE A. foetidus lysophospholipase protein.
XX
KW Lysophospholipase; LPL; degum; vegetable oil; clarifying; hydrolysis;
KW starch hydrolysate; dough.
XX

Query Match 91.7%; Score 22; DB 2; Length 297;
Best Local Similarity 80.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 190 GYSVE 194

RESULT 67
AAW73004
ID AAW73004 standard; protein; 297 AA.
XX
AC AAW73004;
XX
DT 17-OCT-2003 (revised)
DT 27-JAN-1999 (first entry)
XX
DE Aspergillus tubigenensis Lip A enzyme.
XX
KW Lip A; lipase; enzyme; recombinant; filamentous fungus; improver;
KW breadmaking dough; gluten; bread; mechanical strength; sliceability;
KW Aspergillus; baked product; lipA gene; fermentation.
XX
OS Aspergillus tubigenensis.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT Peptide /note= "signal peptide"
FT Protein 28..297
FT Peptide /note= "mature protein"
FT Peptide 28..52
FT Peptide /note= "N-terminal fragment of mature protein; claimed in
FT claim 2"
FT Peptide 136..142
FT Peptide /note= "internal peptide 1, claimed in claim 2"
FT Peptide 148..156
FT Peptide /note= "internal peptide 2, claimed in claim 2"
```

```
OS Aspergillus foetidus.
XX
FN DE19701348-A1.
XX
PD 23-JUL-1998.
XX
PF 16-JAN-1997; 97DE-01001348.
XX
PR 16-JAN-1997; 97DE-01001348.
XX
PA (ROHG ) ROEHM GMBH.
XX
PI Loeffler F, Jungschaffer G, Khanh QN, Schuster E, Sproessler B;
PI Wolf S;
XX
DR WPI; 1998-399863/35.
DR N-PSDB; AAV46288.
XX
PT New two-chain form of Aspergillus lyso:phospholipase - useful for
PT degumming plant oils, as baking aid and generally for hydrolysis of
PT phospholipase.
XX
PS Disclosure; Page 13-14; 16pp; German.
XX
CC This sequence represents a lysophospholipase (LPL) from Aspergillus
CC foetidus. In the event of cleavage of this protein, the fragments either
CC remain connected by at least one bond cleavable under reducing
CC conditions, or at least one fragment has LPLase activity. This protein is
CC used to degum vegetable oils e.g. as a baking aid (to improve workability
CC of dough) and for clarifying starch hydrolysate or generally to hydrolyse
CC phospholipids. The protein can be produced inexpensively, in high purity
CC and in large quantity
XX
SQ Sequence 297 AA;

Query Match 91.7%; Score 22; DB 2; Length 297;
Best Local Similarity 80.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 190 GYSVE 194

RESULT 67
AAW73004
ID AAW73004 standard; protein; 297 AA.
XX
AC AAW73004;
XX
DT 17-OCT-2003 (revised)
DT 27-JAN-1999 (first entry)
XX
DE Aspergillus tubigenensis Lip A enzyme.
XX
KW Lip A; lipase; enzyme; recombinant; filamentous fungus; improver;
KW breadmaking dough; gluten; bread; mechanical strength; sliceability;
KW Aspergillus; baked product; lipA gene; fermentation.
XX
OS Aspergillus tubigenensis.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT Peptide /note= "signal peptide"
FT Protein 28..297
FT Peptide /note= "mature protein"
FT Peptide 28..52
FT Peptide /note= "N-terminal fragment of mature protein; claimed in
FT claim 2"
FT Peptide 136..142
FT Peptide /note= "internal peptide 1, claimed in claim 2"
FT Peptide 148..156
FT Peptide /note= "internal peptide 2, claimed in claim 2"
```

XX PN W09845453-A1.  
XX PD 15-OCT-1998.  
XX PF 03-APR-1998; 98WO-DK000137.  
XX PR 09-APR-1997; 97DK-00000400.  
XX PA (DANI-) DANISCO AS.  
XX PI Poulsen CH, Soe JB, Rasmussen P, Madrid SM, Zargahi MR;  
XX DR WPI; 1998-568355/48.  
XX DR N-PSDB; AAV07890.  
XX PT New Aspergillus tubigenis lipase - used as dough additive to reduce pore  
XX PT size in the crumb, increase pore homogeneity and increase gluten index.  
XX PS Claim 11; Page 36-37; 95pp; English.  
XX CC This sequence represents the Aspergillus tubigenis LipA enzyme, having  
CC lipase activity. This new lipase enzyme retains at least 80 percentage  
CC activity after 4 days at 20 deg. C and pH 3.5-8; retains at least 60  
CC (especially 80) percentage activity after 1 hour at 60 deg. C in 0.1 M  
CC sodium acetate buffer, and has isoelectric point 3.5-4.5 (preferably 4.0  
CC -4.2) as measured by isoelectric focusing. Cells containing the lipase  
CC encoding lipA gene can be used to produce the enzyme recombinantly. The  
CC new lipase is an improver for breadmaking dough that reduces average pore  
CC diameter, and increases pore homogeneity, in the crumb, and increases the  
CC gluten index (i.e. improves stability of the gluten network). Bread  
CC produced from doughs containing lip A have greater mechanical strength,  
CC especially sliceability and resistance to physical handling, and the  
CC improved gluten network increases tolerance to variations in fermentation  
CC times. (Updated on 17-OCT-2003 to standardise OS field)  
XX SQ Sequence 297 AA;  
Query Match 91.7%; Score 22; DB 2; Length 297;  
Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
Db |||||  
190 GYSVE 194  
RESULT 68  
AAE34500  
ID AAE34500 standard; protein; 297 AA.  
AC AAE34500;  
XX 23-OCT-2003 (revised)  
DT 14-MAY-2003 (first entry)  
XX Aspergillus tubigenis lipase.  
DE  
XX Flour dough; baked product; noodle product; pasta product; cake; lipase;  
KW enzyme; EC 3.1.1.3.  
XX Aspergillus tubigenis.  
OS  
XX W0200294123-A2.  
PN  
XX 28-NOV-2002.  
PD  
XX 17-MAY-2002; 2002WO-IB002792.  
PF  
XX 18-MAY-2001; 2001GB-00012226.  
PR 09-JAN-2002; 2002US-0347007P.  
XX  
XX (DANI-) DANISCO AS.  
PA

XX PI Bojsen K, Poulsen CH, Soe JB;  
XX DR WPI; 2003-120738/11.  
XX DR N-PSDB; AAD52801.  
XX PT Preparing flour dough for preparing baked products, by adding an enzyme  
XX PT that hydrolyses glycolipid and phospholipid, but not triglyceride and/or  
XX PT 1-monoglyceride, to dough components and mixing dough components.  
XX PS Disclosure; Page 104; 107pp; English.  
XX CC The invention relates to a method of preparing flour dough for preparing  
XX CC baked products. The method involves adding an enzyme that hydrolyses  
XX CC glycolipid and phospholipid, but not triglyceride and/or 1-monoglyceride,  
XX CC to dough components and mixing dough components. The method is useful for  
XX CC preparing a flour dough which is useful for producing baked products,  
XX CC noodle products, pasta products and cakes. The invention also provides a  
XX CC method for improving the strength and machinability of doughs and the  
XX CC volume, softness and crumb structure of bread and other baked products.  
XX CC The present sequence is Aspergillus tubigenis lipase (EC 3.1.1.3).  
XX CC (Updated on 23-OCT-2003 to standardise OS field)  
XX SQ Sequence 297 AA;  
Query Match 91.7%; Score 22; DB 6; Length 297;  
Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
Db |||||  
190 GYSVE 194  
RESULT 69  
ADF82797  
ID ADF82797 standard; protein; 297 AA.  
XX ADF82797;  
XX 26-FEB-2004 (first entry)  
XX Aspergillus niger phospholipase PLP06.  
DE Phospholipase; enzyme.  
KW Aspergillus niger.  
OS  
XX W02003097825-A2.  
PN  
XX 27-NOV-2003.  
PD  
XX 21-MAY-2003; 2003WO-EP005450.  
PF  
XX 21-MAY-2002; 2002EP-00100521.  
PR 21-MAY-2002; 2002EP-00100524.  
XX 21-MAY-2002; 2002EP-00100528.  
PR 21-MAY-2002; 2002EP-00100538.  
XX (STAM ) DSM IP ASSETS BV.  
PA  
XX Albermann K, Kemmer W, Kimpel E, Maier D, Spreafico F, Stock A;  
XX PI Wagner C, Boer DL, Meima RB;  
XX WPI; 2004-022875/02.  
XX DR N-PSDB; ADF82795, ADF82796.  
XX New polynucleotide encoding a phospholipase, useful in producing dough or  
XX PT baked product of dough.  
XX Claim 13; SEQ ID NO 6; 88pp; English.  
XX The present sequence is the protein sequence of Aspergillus niger  
CC

CC phospholipase PLP06. This enzyme has a calculated molecular weight of  
 CC 31.694. In an example from the invention, an *A. niger* culture  
 CC ultrafiltrate containing PLP06 had amylase (5.4 U/mg protein),  
 CC phospholipase A (67.3 U/mg), phospholipase C (0.0 U/mg),  
 CC lysophospholipase (50.0 U/mg) and galactolipase (124.5 U/mg) activities.  
 CC The invention provides *A. niger* phospholipase polynucleotides, vectors  
 CC and recombinant host cells, as well as phospholipase polypeptides and  
 CC methods for their recombinant production. The phospholipases are useful  
 CC for the production of dough and baked products (claimed), and can also be  
 CC used in industrial processes and in the diagnosis of fungal infections.  
 XX  
 SQ Sequence 297 AA;

Query Match 91.7%; Score 22; DB 8; Length 297;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 190 GYSVE 194

RESULT 70  
 ADN59253  
 ID ADN59253 standard; protein; 297 AA.

XX AC ADN59253;

XX DT 01-JUL-2004 (first entry)

XX DE Aspergillus tubingensis lipase A.

XX KW lipase; galactosyl diglyceride hydrolysis; flour dough; monoglyceride;  
 KW glycolipid hydrolysis; monogalactosyl diglyceride;  
 KW digalactosyl diglyceride; galactosyl monoglyceride; pore homogeneity;  
 KW pore diameter; dough; lipid; gluten network; lipase A.

XX OS Aspergillus tubingensis.

XX FN US2004071853-A1.

XX PD 15-APR-2004.

XX PF 16-JUN-2003; 2003US-00462527.

XX PR 09-APR-1997; 97DK-00000400.

XX PR 03-APR-1998; 98WO-DK000136.

XX PR 09-JAN-2002; 2002US-00040394.

XX PA (SOEJ/) SOE J B.

XX PA (POUL/) POULSEN C H.

XX PA (RASM/) RASMUSSEN P.

XX PA (MADR/) MADRID S M.

XX PA (ZARG/) ZARGAHI M R.

XX PI Soe JB, Poulsen CH, Rasmussen P, Madrid SM, Zargahi MR;

XX WPI: 2004-328553/30.

XX DR N-PSDB; ADN59262.

XX PS Claim 23; SEQ ID NO 9; 38pp; English.

XX CC The invention describes a polypeptide (I) having lipase activity, where  
 CC (I) is capable of hydrolysing at least 10% of galactosyl diglycerides  
 CC normally present in a flour dough to monoglycerides, where (I) retains at  
 CC least 82.5% activity after 4 days at room temperature and at a pH of 3.5-  
 CC 8, and where the polypeptide is capable of hydrolysing glycolipids  
 CC (monogalactosyl diglyceride and digalactosyl diglyceride) that are

CC normally present in a flour to the corresponding galactosyl  
 CC monoglycerides. (I) is useful for preparing a baked product having  
 CC improved pore homogeneity and reduced average pore diameter. The dough  
 CC does not contain added lipids. The polypeptide is added to the dough in  
 CC an amount that is in the range of 5000-30000 lipase units (IUS) per kg  
 CC flour. (I) is useful for improving the stability of a gluten network in a  
 CC dough, imparting improved pore homogeneity, reducing pore diameter of a  
 CC baked product made from the dough or its combination that involves adding  
 CC (I) to the dough, the gluten index being determined by Glutomatic 2200  
 CC apparatus. This is the amino acid sequence of *Aspergillus tubingensis*  
 CC lipase A.

XX SQ Sequence 297 AA;

Query Match 91.7%; Score 22; DB 8; Length 297;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 190 GYSVE 194

RESULT 71  
 ABG23387

ID ABG23387 standard; protein; 298 AA.

XX AC ABG23387;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #23378.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX DR N-PSDB; AAS87574.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 53746; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 CC SQ Sequence 298 AA;

Query Match 91.7%; Score 22; DB 4; Length 298;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

DB 228 GYSVE 232

RESULT 72  
 ABP38577  
 ID ABP38577 standard; protein; 300 AA.

AC ABP38577;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3422.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
 KW antibacterial; gene therapy.

XX Staphylococcus epidermidis.

PN US6380370-B1.

XX 30-APR-2002.

PF 13-AUG-1998; 98US-00134001.

PR 14-AUG-1997; 97US-0055779P.

PR 08-NOV-1997; 97US-0064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Doucette-Stamm LA, Bush D;

DR WPI: 2002-381255/41.

DR N-PSDB; ABN91122.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermis

PT polypeptide, useful for diagnosing and treating bacterial infections.

XX Disclosure; SEQ ID NO 3422; 267pp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have  
 CC antibacterial activity and can be used in gene therapy. The sequences can  
 CC also be used in the diagnosis and treatment of bacterial infections,  
 CC particularly S. epidermidis infections. The sequences can be used to  
 CC screen for compounds able to interfere with the S. epidermidis life cycle  
 CC or inhibit S. epidermidis infection. N.B. The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the USPTO web site

XX Sequence 300 AA;

Query Match 91.7%; Score 22; DB 5; Length 300;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 DB 203 GYSVE 207  
 RESULT 73  
 ABB07518  
 ID ABB07518 standard; protein; 302 AA.

XX ABB07518;

DT 23-APR-2002 (first entry)

DE Human drug metabolizing enzyme (DME) (ID: 7474438CD1).

XX Drug metabolizing enzyme; DME; anti-allergic; antianemic; antiasthmatic;  
 KW osteopathic; antineumatic; antiarthritic; dermatological; nephrotropic;  
 KW anti-inflammatory; vulnery; antibacterial; virucide; antiparasitic;  
 KW protozoacide; fungicide; antihelminthic; cytostatic; ophthalmological;  
 KW antiarteriosclerotic; hepatotropic; antidiabetic; anorectic; human;  
 KW thrombolytic; metabolic; anticoagulant; antithyroid; gynecological;  
 KW antiangular; antiulcer; antiarrhoeic; laxative; enzyme.

XX Homo sapiens.

XX WO200204612-A2.

XX 17-JAN-2002.

XX 05-JUL-2001; 2001WO-US021324.

XX 07-JUL-2000; 2000US-0216804P.

XX 14-JUL-2000; 2000US-0218948P.

XX 21-JUL-2000; 2000US-0220037P.

XX 28-JUL-2000; 2000US-0221837P.

XX (INCY-) INCYTE GENOMICS INC.

XX Baughn MR, Bruns CM, Das D, Ding L, Elliott VS, Gandhi AR;  
 PI Hafalia AJA, Kearney L, Khan FA, Lal P, Lee EA, Lu DAM, Lu Y;  
 PI Nguyen DB, Patterson C, Ramkumar J, Ring HZ, Sanjanwala MS, Tang YT;  
 PI Thangavelu K, Thornton M, Tribouley CM, Walia NK, Xu Y, Yang J;  
 PI Yao MG, Yue H;

XX WPI: 2002-164641/21.

XX N-PSDB; ABA94721.

XX Novel drug metabolizing enzymes and polynucleotides encoding the enzymes,  
 PT useful for treating, diagnosing or preventing autoimmune/inflammatory,  
 PT cell proliferative, developmental and endocrine disorders.

XX Claim 1; Page 142-143; 167pp; English.

XX The invention provides human drug metabolizing enzyme (DME) polypeptides  
 CC and polynucleotides. DMEs can be expressed by standard recombinant  
 CC methodology. The DME polypeptides, polynucleotides, modulators and  
 CC antibodies are useful from diagnosing, treating and preventing autoimmune  
 CC /inflammatory (e.g. allergies, anemia, asthma, osteoporosis, rheumatoid  
 CC arthritis, atopic dermatitis, glomerulonephritis and irritable bowel  
 CC syndrome, trauma, and bacterial, viral, parasitic, protozoal, fungal or  
 CC helminthic infections), cell proliferative (e.g. cancer, atherosclerosis,  
 CC and hepatitis), developmental (e.g. cerebral palsy and cataract),  
 CC endocrine (e.g. thrombosis, goiter, hyperplasia, amenorrhea and  
 CC gynecomasia), eye (e.g. iritis and glaucoma, metabolic (e.g. Addison's  
 CC disease, diabetes and obesity), and gastrointestinal disorders (e.g.  
 CC anorexia, nausea, abdominal angina, ulcerative colitis, diarrhoea and  
 CC constipation). The DME polypeptides are useful in drug screening  
 CC techniques, and to analyse the proteome of a tissue or cell type. The DME  
 CC polynucleotides are useful for creating knock-in humanized animals or  
 CC transgenic animals to model human diseases, and in somatic or germline  
 CC gene therapy. The present sequence represents a human DME polypeptide  
 CC sequence

XX SQ Sequence 302 AA;  
Query Match 91.7%; Score 22; DB 5; Length 302;  
Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 232 GYSVE 236

RESULT 74  
AAG82514  
ID AAG82514 standard; protein; 307 AA.  
XX AC AAG82514;  
XX DT 03-SEP-2001 (first entry)  
XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:2122.  
XX KW Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;  
XX endocarditis.  
XX OS Staphylococcus epidermidis.  
XX PN WO200134809-A2.  
XX PD 17-MAY-2001.  
XX PF 09-NOV-2000; 2000WO-US030782.  
XX PR 09-NOV-1999; 99US-0164258P.  
XX PA (GLAX ) GLAXO GROUP LTD.  
XX PI Kimmerly WJ;  
XX WPI; 2001-316495/33.  
XX DR N-PSDB; AAH53364.  
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
XX useful for vaccinating against infections, e.g. endocarditis.  
XX PS Claim 18; Page 572; 2188pp; English.  
XX CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
XX (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)  
XX and (II) can have antibacterial activity and therefore can be used in  
XX vaccination. The nucleic acids (I) may be used to produce the S.  
XX epidermidis polypeptides (II) via the production of vectors containing  
XX them which are used to produce hosts cells which express the  
XX polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
XX used to vaccinate subjects and to raise antibodies against the bacteria.  
XX The polypeptides may also be used to assay for other inhibitors of their  
XX activity and therefore identify compounds that may be used for the  
XX treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
XX AAH5090 represent specifically claimed S. epidermidis genomic DNA  
XX polynucleotide sequences from the present invention. AAH55091 to AAH55098  
XX represent oligonucleotide sequences and primers which are used in the  
XX exemplification of the present invention. N.B. The present invention  
XX specifically claims all the polynucleotide sequences given in the  
XX sequence listing of the present specification, however the sequence  
XX listing only goes up to SEQ ID NO:4454 so even though the sequences are given  
XX in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present  
XX for SEQ ID NO:4455 to 4464

SQ Sequence 307 AA;  
Query Match 91.7%; Score 22; DB 4; Length 307;  
Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 256 GYAVE 260

RESULT 75  
ABU43080  
ID ABU43080 standard; protein; 307 AA.  
XX AC ABU43080;  
XX DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #28607.  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Staphylococcus epidermidis.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX DR WPI; 2003-029926/02.  
XX DR N-PSDB; ACA46950.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 25; SEQ ID NO 71004; 1766pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX the 6213 antisense sequences given in the specification where expression  
XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX (1) a vector comprising a promoter operably linked to the nucleic acid  
XX encoding a polypeptide whose expression is inhibited by the antisense  
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX polypeptide or its fragment whose expression is inhibited by the  
XX antisense nucleic acid; (4) an antibody capable of specifically binding  
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX proliferation or the activity of a gene in an operon required for  
XX the gene product or that has an activity against a biological pathway  
XX required for proliferation, or that inhibits cellular proliferation; (8)  
XX identifying a gene required for cellular proliferation or the biological  
XX pathway in which a proliferation-required gene or its gene product lies  
XX or a gene on which the test compound that inhibits proliferation of an  
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX compound's activity; (11) a culture comprising strains in which the gene  
XX product is overexpressed or underexpressed; (12) determining the extent  
XX to which each of the strains is present in a culture or collection of  
XX strains; or (13) identifying the target of a compound that inhibits the  
XX proliferation of an organism. The antisense nucleic acids are useful for  
XX identifying proteins or screening for homologous nucleic acids required  
XX for cellular proliferation to isolate candidate molecules for rational  
XX drug discovery programs, or for screening homologous nucleic acids  
XX required for proliferation in cells other than S. aureus, S. typhimurium,  
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 307 AA;  
Query Match 91.7%; Score 22; DB 6; Length 307;  
Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GYXVE 5  
Db 256 GYAVE 260  
Search completed: November 1, 2004, 21:31:00  
Job time : 158 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:15:40 ; Search time 24 Seconds  
(without alignments)  
24.054 Million cell updates/sec

Title: US-10-030-194A-6

Perfect score: 24

Sequence: 1 GYXVEX 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

PIR 79:\*

1: piri:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	91.7	74	2 JU0240	nuclear matrix pro
2	22	91.7	76	2 A13049	transcription regu
3	22	91.7	78	1 RDECD5	dihydrofolate redu
4	22	91.7	78	1 RDECD6	dihydrofolate redu
5	22	91.7	78	1 RDECD8	dihydrofolate redu
6	22	91.7	78	2 T08517	dihydrofolate redu
7	22	91.7	78	2 S32183	dihydrofolate redu
8	22	91.7	82	2 S58349	H+-transporting tw
9	22	91.7	82	2 T07211	H+-transporting tw
10	22	91.7	86	2 C97083	uncharacterized pr
11	22	91.7	97	2 C37202	hypothetical prote
12	22	91.7	108	2 S64602	hypothetical prote
13	22	91.7	109	2 G71077	hypothetical prote
14	22	91.7	119	2 B98236	exs1 protein prote
15	22	91.7	120	2 A69453	hypothetical prote
16	22	91.7	122	2 B87323	chemotaxis protein
17	22	91.7	123	2 C96604	hypothetical prote
18	22	91.7	131	2 D84401	30S ribosomal prot
19	22	91.7	136	2 H83194	hypothetical prote
20	22	91.7	136	2 D75265	conserved hypothet
21	22	91.7	137	2 G59473	hypothetical prote
22	22	91.7	142	2 H72520	hypothetical prote
23	22	91.7	148	2 AF1081	50S ribosomal prot
24	22	91.7	148	2 AG1438	50S ribosomal prot
25	22	91.7	148	2 G75066	hypothetical prote
26	22	91.7	151	2 B71166	hypothetical prote
27	22	91.7	152	2 B64485	hypothetical prote
28	22	91.7	154	2 S19721	pyruvate dehydroge
29	22	91.7	157	2 PH0201	hypothetical prote

30	22	91.7	158	2 D83384	probable transcrip
31	22	91.7	162	2 H69504	hypothetical prote
32	22	91.7	168	2 T35779	probable secreted
33	22	91.7	174	2 H70702	hypothetical prote
34	22	91.7	174	2 A70395	hypothetical prote
35	22	91.7	179	2 B71274	probable translati
36	22	91.7	179	2 B72459	hypothetical prote
37	22	91.7	180	2 T34851	probable secreted
38	22	91.7	180	2 T06718	hypothetical prote
39	22	91.7	182	2 E87289	conserved hypothet
40	22	91.7	187	2 AG2528	hypothetical prote
41	22	91.7	188	2 D87705	intracellular sept
42	22	91.7	208	2 AF1504	amidotransferase
43	22	91.7	208	2 A41030	tropoin I, cardia
44	22	91.7	209	2 AH2716	NTP pyrophosphohy
45	22	91.7	209	2 A54984	ELF-1 protein prec
46	22	91.7	211	2 T32354	hypothetical prote
47	22	91.7	213	2 JE0322	ephra-A2 - human
48	22	91.7	214	2 J02001	hypothetical 24K p
49	22	91.7	220	2 A72153	Q2L protein - vari
50	22	91.7	220	2 H42504	M2L protein - vacc
51	22	91.7	220	2 T28457	hypothetical prote
52	22	91.7	220	2 I36838	O2L protein - vari
53	22	91.7	220	2 C84383	hypothetical prote
54	22	91.7	228	2 P86930	probable two-compo
55	22	91.7	229	2 H64367	dolichyl-phosphate
56	22	91.7	230	2 A70821	hypothetical prote
57	22	91.7	231	2 S45089	hypothetical prote
58	22	91.7	233	2 G75075	hypothetical prote
59	22	91.7	236	2 AF2051	two-component svst
60	22	91.7	238	2 B84167	glucose-1-phosphat
61	22	91.7	243	2 C82393	probable amino aci
62	22	91.7	244	2 G75259	conserved hypothet
63	22	91.7	253	2 T45446	probable two-compo
64	22	91.7	254	2 C97498	hypothetical 21.7K
65	22	91.7	258	2 G84279	stationary phase s
66	22	91.7	270	2 F70432	conserved hypothet
67	22	91.7	272	2 B95852	conserved hypothet
68	22	91.7	280	1 WMBP8H	gene 8.5 protein -
69	22	91.7	280	1 WMBP8H	HE65 orf105 - Bomb
70	22	91.7	289	2 T41846	hypothetical prote
71	22	91.7	293	2 T27430	arginase (EC 3.5.3
72	22	91.7	296	1 S55795	probable binding p
73	22	91.7	299	2 D83010	glycerol-3-phospha
74	22	91.7	313	2 A70441	probable ATP-bind
75	22	91.7	323	2 D72484	cathepsin - Orgyia
76	22	91.7	324	2 T10394	cathepsin - Choris
77	22	91.7	324	2 S62735	hypothetical prote
78	22	91.7	325	2 C89879	hypothetical prote
79	22	91.7	334	2 A12110	hypothetical prote
80	22	91.7	334	2 G83533	probable polyferre
81	22	91.7	334	2 G75344	anthranilate phosp
82	22	91.7	337	1 NPBS	phenylalanine-trNA
83	22	91.7	344	1 YFBSA	probable acetylpol
84	22	91.7	345	2 B83605	GRP-binding regula
85	22	91.7	345	2 A40990	spermidine/putresc
86	22	91.7	345	2 G69450	acetylpolylamine am
87	22	91.7	346	2 H83469	precorrin-3 methyl
88	22	91.7	351	2 H69053	acetylpolylamine am
89	22	91.7	359	1 B59266	citrate synthase (
90	22	91.7	361	2 D87700	carbamoyl-phosphat
91	22	91.7	363	2 AD1304	hypothetical prote
92	22	91.7	363	2 AD1676	hypothetical prote
93	22	91.7	363	2 A72702	E2 protein - human
94	22	91.7	365	1 W2MLHS	sepium site-determ
95	22	91.7	366	2 AH2217	GRP-binding protei
96	22	91.7	370	1 Q8HS4C	GRP-binding protei
97	22	91.7	370	2 C84267	E2 protein - human
98	22	91.7	372	1 W2WLJ1	2-hydroxyglutaryl-
99	22	91.7	373	2 G64300	prv43 protein - su
100	22	91.7	373	2 T49806	hypothetical prote
101	22	91.7	373	2 T45043	Ig heavy chain V r
102	22	91.7	374	2 S69339	

103	22	91.7	376	1	E69957	gamma-D-glutamyl-L	176	22	91.7	748	2	T49633	glucan 1,4-alpha-g
104	22	91.7	379	2	T19773	hypothetical prote	177	22	91.7	773	2	B97071	protein containing
105	22	91.7	379	2	T49919	hypothetical prote	178	22	91.7	775	2	T37848	probable cleavage
106	22	91.7	385	1	YKYT	citrate (si)-synth	179	22	91.7	782	2	A82940	hypothetical prote
107	22	91.7	389	1	J50443	alanine racemase (	180	22	91.7	791	2	F81056	hemoglobin recepto
108	22	91.7	389	2	D30498	acyl-CoA dehydroge	181	22	91.7	806	2	AB1015	glycerol-3-phospha
109	22	91.7	390	2	T22810	hypothetical prote	182	22	91.7	810	2	D64090	glycerol-3-phospha
110	22	91.7	394	2	E75104	hypothetical prote	183	22	91.7	811	2	B82365	hypothetical prote
111	22	91.7	395	2	S70912	CMP-N-acetylneuram	184	22	91.7	816	2	C69493	glycerol-3-phospha
112	22	91.7	396	2	T40010	hypothetical prote	185	22	91.7	825	2	AC0039	glycerol-3-phospha
113	22	91.7	398	2	A82243	acetate kinase VC1	186	22	91.7	827	1	XUECAG	glycerol-3-phospha
114	22	91.7	398	2	D30608	nadh-dependent fla	187	22	91.7	827	2	H91256	glycerol-3-phospha
115	22	91.7	400	1	ORX02R	protein kinase (EC	188	22	91.7	833	2	D86097	maltoase ABC transp
116	22	91.7	401	1	ORX52R	protein kinase (EC	189	22	91.7	833	2	H72205	ribonucleoside-dip
117	22	91.7	404	1	ORXU2R	protein kinase (EC	190	22	91.7	852	2	B84001	GCR3 protein - yea
118	22	91.7	405	2	D70529	hypothetical prote	191	22	91.7	858	2	A44919	cation-transportin
119	22	91.7	410	2	T47586	hypothetical prote	192	22	91.7	879	2	AC2542	104K glycoprotein
120	22	91.7	410	2	T50718	hypothetical prote	193	22	91.7	901	1	WMNVTN	hypothetical lilk
121	22	91.7	411	2	G89074	protein K04A8.5 [i	194	22	91.7	965	2	JQ0058	hypothetical prote
122	22	91.7	419	2	T19260	hypothetical prote	195	22	91.7	993	2	A96750	DNA topoisomerase
123	22	91.7	432	2	T36300	probable secreted	196	22	91.7	1021	2	G75403	hypothetical prote
124	22	91.7	432	2	T34154	hypothetical prote	197	22	91.7	1036	2	E84620	hypothetical prote
125	22	91.7	443	2	F89426	protein M162.5 [im	198	22	91.7	1043	2	F97302	hypothetical prote
126	22	91.7	443	2	G72770	hypothetical prote	199	22	91.7	1047	2	G87398	hypothetical prote
127	22	91.7	452	2	G72770	hypothetical prote	200	22	91.7	1050	2	H90316	AcRb/AcrD/AcrF fam
128	22	91.7	457	2	A57506	palindrome-binding	201	22	91.7	1064	2	F39845	probable isoleucyl
129	22	91.7	459	2	D36889	3-isopropylmalate	202	22	91.7	1071	2	AF1380	glycosidase homolo
130	22	91.7	460	2	D86777	probable 3-isoprop	203	22	91.7	1091	2	AC3921	cobalamin biosynth
131	22	91.7	460	2	D86777	hypothetical prote	204	22	91.7	1116	2	D97695	cobN protein homol
132	22	91.7	461	2	D86294	T24D18.8 protein -	205	22	91.7	1116	2	T19736	hypothetical prote
133	22	91.7	462	2	AE1323	3-isopropylmalate	206	22	91.7	1181	2	T50904	Mg protoporphyrin
134	22	91.7	462	2	AF1694	3-isopropylmalate	207	22	91.7	1236	2	A54080	protein-tyrosine-p
135	22	91.7	465	1	T51095	acid phosphatase (	208	22	91.7	1237	2	S55050	cardiac myosin-bin
136	22	91.7	465	2	T51094	acid phosphatase (	209	22	91.7	1274	2	T34357	hypothetical prote
137	22	91.7	466	2	E95262	hypothetical prote	210	22	91.7	1280	2	AE2192	microbial collagen
138	22	91.7	468	1	B84540	L-seryl-tRNA <sup>Sec</sup> se	211	22	91.7	1282	2	AE2192	probable magnesium
139	22	91.7	470	1	T51096	acid phosphatase (	212	22	91.7	1290	2	T31462	DNA-directed DNA p
140	22	91.7	471	2	S05392	hypothetical prote	213	22	91.7	1292	2	T31462	hypothetical prote
141	22	91.7	472	2	A64759	membrane protein y	214	22	91.7	1339	1	S20052	hypothetical prote
142	22	91.7	472	2	B82564	3-isopropylmalate	215	22	91.7	1352	2	G84473	hypothetical prote
143	22	91.7	474	2	H82564	3-isopropylmalate	216	22	91.7	1374	2	A84888	RNA-directed DNA p
144	22	91.7	475	1	YWBO	tryptophan-tRNA li	217	22	91.7	1412	2	T01610	hypothetical prote
145	22	91.7	477	2	H85981	D-alanyl-D-alanine	218	22	91.7	1585	2	T31611	hypothetical prote
146	22	91.7	477	2	E91136	D-alanyl-D-alanine	219	22	91.7	1659	1	OKBYN2	protein kinase GCN
147	22	91.7	477	2	A54535	serine-type D-Ala-	220	22	91.7	1748	1	JN0786	integrin beta-4 ch
148	22	91.7	487	2	T23776	hypothetical prote	221	22	91.7	1787	2	D69195	integrin beta-4 ch
149	22	91.7	488	2	T47890	hypothetical prote	222	22	91.7	1807	2	JC6319	hypothetical prote
150	22	91.7	493	2	T01495	hypothetical prote	223	22	91.7	1875	2	A36429	hypothetical prote
151	22	91.7	500	2	D97302	hypothetical prote	224	22	91.7	1933	2	F72648	glutamate synthase
152	22	91.7	501	2	T44939	glyceraldehyde-3-p	225	22	91.7	2152	2	T45583	pyrimidine synthe
153	22	91.7	525	2	A96183	D-ribulokinase (EC	226	22	91.7	2194	1	JQ1977	tyrosine synthet
154	22	91.7	532	2	A13103	ribitol kinase [im	227	22	91.7	2225	1	A23443	heparin-binding ep
155	22	91.7	533	2	F70551	hypothetical prote	228	22	91.7	2225	1	I38344	histone H-1-WD8P-2
156	22	91.7	538	2	H72539	probable CTP synth	229	21	87.5	23	I61239	uncharacterized pr	
157	22	91.7	543	2	H84724	probable ARI-like	230	21	87.5	24	A47209	histone H1.a, test	
158	22	91.7	546	2	A10807	hypothetical prote	231	21	87.5	43	S24180	hypothetical prote	
159	22	91.7	550	2	H89780	probable decarboxy	232	21	87.5	51	S68989	conserved hypochet	
160	22	91.7	552	2	S16013	indolepyruvate dec	233	21	87.5	54	T10786	hypothetical prote	
161	22	91.7	553	2	F90032	urocanate hydratase	234	21	87.5	54	T06514	hypothetical prote	
162	22	91.7	553	2	B72863	HE65 protein - Aut	235	21	87.5	57	D64502	hypothetical prote	
163	22	91.7	557	2	G64118	glutamine-tRNA lig	236	21	87.5	57	D64502	hypothetical prote	
164	22	91.7	557	2	T27752	hypothetical prote	237	21	87.5	64	NTSRIC	uncharacterized pr	
165	22	91.7	560	2	E70814	probable pdc prote	238	21	87.5	65	S23080	toxin 3 - scorpion	
166	22	91.7	607	2	T45969	mRNA capping enzym	239	21	87.5	66	AB2108	mercuric transport	
167	22	91.7	618	2	A13303	hypothetical prote	240	21	87.5	67	D83719	hypothetical prote	
168	22	91.7	627	2	B70122	glucose-inhibited	241	21	87.5	68	D83719	hypothetical prote	
169	22	91.7	632	2	AG1366	phosphotransferase	242	21	87.5	69	F70706	hypothetical prote	
170	22	91.7	632	2	AH1735	phosphotransferase	243	21	87.5	74	C84182	hypothetical prote	
171	22	91.7	635	2	T01311	NAD ADP-ribosyltra	244	21	87.5	75	D89353	hypothetical prote	
172	22	91.7	646	2	AB2168	two-component hybr	245	21	87.5	79	G97733	hypothetical prote	
173	22	91.7	679	2	F70425	cation transportin	246	21	87.5	80	S23924	histone H1.b, hepa	
174	22	91.7	684	2	B69308	conserved hypochet	247	21	87.5	82	C75026	hydrogenase expres	
175	22	91.7	701	2	AF2336	potassium-dependen	248	21	87.5	82	D69087	probable transcrip	
	22	91.7	724	2	D84377	protein export [im		21	87.5	83	2	B95974	

249	21	87.5	84	2	H69092	conserved hypothet	322	21	87.5	173	2	T41773	LEF-6 orf28 - Bomb
250	21	87.5	85	2	A12111	hypothetical prote	323	21	87.5	174	2	F75097	adenylylsulfate 3-
251	21	87.5	89	2	S24178	histone H1.d, hepa	324	21	87.5	177	2	T01785	protoporphyrin IX
252	21	87.5	89	2	B69032	conserved hypothet	325	21	87.5	179	2	S62019	hypothetical prote
253	21	87.5	90	2	A72699	probable DNA-direc	326	21	87.5	179	2	H87418	hypothetical prote
254	21	87.5	92	2	T29701	hypothetical prote	327	21	87.5	180	2	D70161	ribosomal protein
255	21	87.5	92	2	E97805	hypothetical prote	328	21	87.5	180	2	F72722	hypothetical prote
256	21	87.5	95	2	E64631	hypothetical prote	329	21	87.5	181	2	T13309	hypothetical prote
257	21	87.5	95	2	C71883	hypothetical prote	330	21	87.5	185	2	S76991	hypothetical prote
258	21	87.5	98	2	F83557	hypothetical prote	331	21	87.5	185	2	AF2073	hypothetical prote
259	21	87.5	99	2	A82651	hypothetical prote	332	21	87.5	186	2	G72590	probable adenylyls
260	21	87.5	103	1	R58B2B	ribosomal protein	333	21	87.5	188	2	AG2484	hypothetical prote
261	21	87.5	103	2	T44394	ribosomal protein	334	21	87.5	190	2	E82984	xanthine phosphori
262	21	87.5	105	1	HSB011	histone H1.1 - bov	335	21	87.5	191	2	A72536	hypothetical prote
263	21	87.5	106	1	H69227	ATP synthase, subu	336	21	87.5	192	2	F81035	modulator of drug
264	21	87.5	106	2	S35258	pyruvate decarboxy	337	21	87.5	192	2	D81979	hypothetical prote
265	21	87.5	111	2	B40508	hypothetical prote	338	21	87.5	192	2	E70414	hypothetical prote
266	21	87.5	115	2	A71088	hypothetical prote	339	21	87.5	194	1	HSTR1	histone H1 - trout
267	21	87.5	115	2	A71088	hypothetical prote	340	21	87.5	194	2	S51309	probable hypoxanth
268	21	87.5	115	2	A71088	hypothetical prote	341	21	87.5	196	2	C72071	DNA-3-methyladenin
269	21	87.5	116	2	C81452	hypothetical prote	342	21	87.5	196	2	E86553	3-methyladenine DN
270	21	87.5	116	2	A81087	hypothetical prote	343	21	87.5	196	2	D70082	DNA-3-methyladenin
271	21	87.5	117	2	C90167	hypothetical prote	344	21	87.5	196	2	F72776	hypothetical prote
272	21	87.5	121	2	E90019	30S ribosomal prot	345	21	87.5	198	2	B83839	xanthine phosphori
273	21	87.5	121	2	F83940	hypothetical prote	346	21	87.5	198	2	AC1384	protein gp51 [Bact
274	21	87.5	124	2	C82805	DNA-binding protei	347	21	87.5	198	2	S77091	hypothetical prote
275	21	87.5	125	2	D82380	chemotaxis protein	348	21	87.5	198	2	B81389	hypothetical prote
276	21	87.5	126	2	B95111	lactoylglutathione	349	21	87.5	200	2	G95219	HAM1 protein [impo
277	21	87.5	128	2	D86809	ferric uptake regu	350	21	87.5	201	2	JC5476	junction-specific
278	21	87.5	128	2	B97170	lactoylglutathione	351	21	87.5	201	2	B72739	hypothetical prote
279	21	87.5	129	2	G95944	probable translati	352	21	87.5	202	2	G90034	hypothetical prote
280	21	87.5	131	2	G83557	lactoylglutathione	353	21	87.5	202	2	S35259	pyruvate decarboxy
281	21	87.5	136	2	C82848	50S ribosomal prot	354	21	87.5	204	2	H72256	conserved hypothet
282	21	87.5	137	2	T45079	hypothetical prote	355	21	87.5	204	2	G87495	hypothetical prote
283	21	87.5	137	2	JH0433	transformation com	356	21	87.5	205	2	B49910	hypothetical prote
284	21	87.5	140	2	B93323	hypothetical prote	357	21	87.5	205	2	C87370	papithyl-prolyl ci
285	21	87.5	143	2	H97979	lactoylglutathione	358	21	87.5	205	2	D89371	DNA polymerase III
286	21	87.5	145	2	G81263	probable periplasm	359	21	87.5	206	1	HSTR1R	histone H1 - rainb
287	21	87.5	146	2	E90556	hypothetical prote	360	21	87.5	206	2	C83590	probable transcrip
288	21	87.5	148	2	B95281	hypothetical prote	361	21	87.5	206	2	JC7320	K562 cell-derived
289	21	87.5	149	2	A69220	conserved hypothet	362	21	87.5	207	2	S39103	ubiquinol-cytochro
290	21	87.5	149	2	H75047	Ni, Fe-Hydrogenase	363	21	87.5	207	2	A56190	titin - rat (fragm
291	21	87.5	151	2	C71113	Ni, Fe-Hydrogenase	364	21	87.5	207	2	JH0550	histone H1t - huma
292	21	87.5	152	2	S74125	cysteine proteinas	365	21	87.5	207	2	AH1548	3-methyladenine DN
293	21	87.5	152	2	E95256	transcription regu	366	21	87.5	208	1	A38432	heparin-binding EG
294	21	87.5	152	2	AF1103	transcription repr	367	21	87.5	208	1	A41914	diphtheria toxin re
295	21	87.5	152	2	AF1465	transcription repr	368	21	87.5	208	1	HSRT1T	histone H1t - rat
296	21	87.5	152	2	T16760	hypothetical prote	369	21	87.5	208	2	C64084	drug activity modu
297	21	87.5	153	2	G95988	conserved hypothet	370	21	87.5	208	2	AF1145	amidotransferases
298	21	87.5	154	2	S66112	hypothetical prote	371	21	87.5	208	2	I70195	histone H1t - rhes
299	21	87.5	154	2	B83304	hypothetical prote	372	21	87.5	208	2	S43434	histone H1, testic
300	21	87.5	154	2	T39892	probable prefoldin	373	21	87.5	208	2	G72408	conserved hypothet
301	21	87.5	155	1	NRECH	ribonuclease H [EC	374	21	87.5	209	2	AB1444	gp51 [Bacteriophag
302	21	87.5	155	2	H81884	hypothetical prote	375	21	87.5	209	2	D97056	competence ComEA p
303	21	87.5	155	2	H69285	adenylylsulfate 3-	376	21	87.5	211	1	HSPG1T	histone H1t - pig
304	21	87.5	155	2	B95506	ribonuclease H [EC	377	21	87.5	212	2	A28470	histone H1 - mouse
305	21	87.5	155	2	B90655	ribonuclease H [EC	378	21	87.5	212	2	A37471	hypothetical orf3
306	21	87.5	155	2	T13308	hypothetical prote	379	21	87.5	212	2	G90353	2-halocalkanoic aci
307	21	87.5	156	2	D83662	transcription repr	380	21	87.5	213	1	HSHU11	histone H1-2 [vali
308	21	87.5	157	2	B81137	modulator of drug	381	21	87.5	213	1	HSTRB13	histone H1.3 - rab
309	21	87.5	157	2	D75151	hypothetical prote	382	21	87.5	213	2	S43949	histone H1 - mouse
310	21	87.5	160	2	C72501	hypothetical prote	383	21	87.5	213	2	RGBSA	regulatory protein
311	21	87.5	161	2	AG1132	hypothetical prote	384	21	87.5	214	1	RGBSA	histone H1.1 - hum
312	21	87.5	162	2	E87351	hypothetical prote	385	21	87.5	215	2	S26363	histone H1.1 - hum
313	21	87.5	163	2	S49633	hypothetical prote	386	21	87.5	215	2	B84078	hypothetical prote
314	21	87.5	164	2	G97079	PTS system, fructo	387	21	87.5	215	2	E75058	hypothetical prote
315	21	87.5	168	2	T13326	hypothetical prote	388	21	87.5	216	2	S59942	GTP-binding protei
316	21	87.5	168	2	T13631	hypothetical prote	389	21	87.5	217	2	JH0159	histone H1d - rat
317	21	87.5	169	2	D81180	16S rRNA processin	390	21	87.5	217	2	F70817	probable ABC trans
318	21	87.5	169	2	B81924	probable 16S rRNA	391	21	87.5	218	1	HSC11	histone H1.02 - ch
319	21	87.5	172	2	AG2149	hypothetical prote	392	21	87.5	218	2	A23055	histone H1 - musco
320	21	87.5	173	2	F98121	transcription regu	393	21	87.5	219	1	HSHU1B	histone H1.01 - ch
321	21	87.5	173	2	D72853	late expression fa	394	21	87.5	219	2	C28456	histone H1.11R - c

395	21	87.5	219	2	I49742	histone H1 - mouse	468	21	87.5	274	2	E81274	probable flagellar
396	21	87.5	220	2	A28456	histone H1.10 - ch	469	21	87.5	275	2	E82179	ABC transporter, p
397	21	87.5	221	1	B40335	histone H1-3 [vali	470	21	87.5	276	2	S28749	NADH2 dehydrogen
398	21	87.5	221	1	S33219	histone H1-C - Afr	471	21	87.5	275	2	D81947	probable amino aci
399	21	87.5	221	2	S49482	histone H1 - mouse	472	21	87.5	275	2	H81158	amino acid ABC tra
400	21	87.5	221	2	G72665	hypothetical prote	473	21	87.5	275	2	S75190	hypothetical prote
401	21	87.5	223	1	YLHUP	serum amyloid P-co	474	21	87.5	277	2	AD3548	creatininase [EC 3
402	21	87.5	223	2	S49492	histone H1 - mouse	475	21	87.5	278	2	D71128	hypothetical prote
403	21	87.5	224	2	D28456	histone H1.03 - ch	476	21	87.5	279	2	S03804	hypothetical prote
404	21	87.5	224	2	D90036	hypothetical prote	477	21	87.5	280	2	C69343	hypothetical prote
405	21	87.5	225	2	B28456	histone H1.11L - c	478	21	87.5	281	2	S73895	diaminopimelate ep
406	21	87.5	225	2	S51660	histone H1-5 [vali	479	21	87.5	284	2	B84118	hypothetical prote
407	21	87.5	226	1	T20447	hypothetical prote	480	21	87.5	285	2	H64213	hypothetical prote
408	21	87.5	226	2	D83875	hypothetical prote	481	21	87.5	285	2	B83047	conserved hypotet
409	21	87.5	229	2	C69971	conserved hypotet	482	21	87.5	285	2	H69802	prephenate dehydro
410	21	87.5	230	2	A87730	protein Y23H5A.1 [	483	21	87.5	286	2	B97010	amino acid ABC tra
411	21	87.5	231	2	A93022	hypothetical prote	484	21	87.5	287	2	D69760	ribokinase homolog
412	21	87.5	231	2	G95008	phosphorylase, Pnp	485	21	87.5	287	2	T46874	peptide chain rele
413	21	87.5	234	2	G95019	conserved hypotet	486	21	87.5	288	2	E84100	hypothetical prote
414	21	87.5	235	1	A59036	cytochrome C554, t	487	21	87.5	288	2	T21732	NADH2 dehydrogen
415	21	87.5	239	2	A33506	transcription regu	488	21	87.5	290	2	S26016	hypothetical prote
416	21	87.5	240	2	H69533	hypothetical prote	489	21	87.5	290	2	T24926	vegetative storage
417	21	87.5	241	2	A49903	phosphate regulato	490	21	87.5	291	2	T08848	hypothetical prote
418	21	87.5	241	2	AC2581	two component resp	491	21	87.5	291	2	A83334	hypothetical prote
419	21	87.5	242	2	AH1180	amino acid ABC tra	492	21	87.5	292	2	S01650	probable signal pe
420	21	87.5	242	2	A11537	amino acid ABC tra	493	21	87.5	292	2	G81944	hydroxymethylbilan
421	21	87.5	242	2	T50297	probable aminopept	494	21	87.5	292	2	A64371	conserved hypotet
422	21	87.5	244	1	UESY25	vegetative storage	495	21	87.5	292	2	G90215	NADH2 dehydrogen
423	21	87.5	245	1	NDEC95	type II site-speci	496	21	87.5	294	2	T11045	dioxigenase [impor
424	21	87.5	245	2	F96913	glutamine ABC tran	497	21	87.5	295	2	AD2887	hypothetical prote
425	21	87.5	247	2	F71545	probable oligopept	498	21	87.5	295	2	A97663	hypothetical prote
426	21	87.5	247	2	T20027	hypothetical prote	499	21	87.5	296	2	T13885	NADH2 dehydrogen
427	21	87.5	249	2	T06441	storage protein ho	500	21	87.5	296	2	AG3388	exodeoxyribonuclea
428	21	87.5	250	2	T43153	probable aldehyde	501	21	87.5	296	2	AB0557	cytochrome o ubiq
429	21	87.5	250	2	D83549	hypothetical prote	502	21	87.5	296	2	S51746	hypothetical prote
430	21	87.5	252	2	C83837	hypothetical prote	503	21	87.5	296	2	H95395	protein [imported
431	21	87.5	253	2	B97016	imidazoleglycerol-	504	21	87.5	298	2	AH0947	probable sugar kin
432	21	87.5	254	1	UESY27	vegetative storage	505	21	87.5	298	2	AG3165	conserved hypotet
433	21	87.5	254	2	S08511	DNA-binding respon	506	21	87.5	298	2	B86770	UDP-N-acetylenolp
434	21	87.5	254	2	B87396	conserved hypotet	507	21	87.5	299	2	S60971	probable membrane
435	21	87.5	254	2	B72374	conserved hypotet	508	21	87.5	300	2	T27158	hypothetical prote
436	21	87.5	254	2	B72880	hypothetical prote	509	21	87.5	300	2	G87999	protein Y54E5A.1 [
437	21	87.5	255	2	T75222	hypothetical prote	510	21	87.5	300	2	H82955	hypothetical prote
438	21	87.5	255	2	T35217	hypothetical prote	511	21	87.5	300	2	AG2955	hypothetical prote
439	21	87.5	255	2	T22863	hypothetical prote	512	21	87.5	300	2	E98327	SEC14 protein - Ye
440	21	87.5	256	2	D64637	amino acid ABC tra	513	21	87.5	301	1	A37766	SEC14 protein - Ye
441	21	87.5	257	2	E96780	thauartin-like pro	514	21	87.5	301	2	S57923	caf1R protein - Ye
442	21	87.5	257	2	E71877	probable amino aci	515	21	87.5	301	2	S19097	caf1R protein - Ye
443	21	87.5	257	2	H84713	hypothetical prote	516	21	87.5	301	2	T14705	hypothetical prote
444	21	87.5	257	2	T36268	probable DNA-bindi	517	21	87.5	302	2	H84184	acyl carrier prote
445	21	87.5	258	2	T20205	hypothetical prote	518	21	87.5	302	2	H69551	proline/glycine be
446	21	87.5	259	2	T12407	NADH2 dehydrogen	519	21	87.5	303	2	G97081	stearyl-CoA 9-des
447	21	87.5	263	2	A54060	nicotinamide N-met	520	21	87.5	305	2	T52111	dihydroorotase deh
448	21	87.5	264	2	A75057	hypothetical prote	521	21	87.5	306	2	B70304	branched-chain ami
449	21	87.5	264	2	A82295	sporulation protei	522	21	87.5	307	2	B36125	histone deacetylase
450	21	87.5	264	2	A82295	conserved hypotet	523	21	87.5	307	2	T15012	fl capsule positiv
451	21	87.5	265	2	A97363	transcription regu	524	21	87.5	307	2	T04141	hypothetical prote
452	21	87.5	265	2	T34399	hypothetical prote	525	21	87.5	307	2	A75020	hypothetical prote
453	21	87.5	266	2	A98177	hypothetical prote	526	21	87.5	307	2	T48166	hypothetical prote
454	21	87.5	266	2	AC3110	creatinine amidohy	527	21	87.5	308	2	D89915	conserved hypotet
455	21	87.5	267	2	A97762	hypothetical prote	528	21	87.5	310	1	B69475	probable sugar kin
456	21	87.5	268	2	E69761	probable glutamine	529	21	87.5	310	2	C72752	hypothetical prote
457	21	87.5	268	2	AD2377	ATP-binding protei	530	21	87.5	311	2	S77803	mevalonate kinase
458	21	87.5	268	2	T50354	5-amino-6-(5-phosp	531	21	87.5	311	2	F64435	hypothetical prote
459	21	87.5	268	2	AF1368	amino acid ABC tra	532	21	87.5	312	2	C71136	hypothetical prote
460	21	87.5	269	2	AF1737	amino acid ABC tra	533	21	87.5	312	2	H71438	hypothetical prote
461	21	87.5	269	2	T74420	phosphate transpor	534	21	87.5	315	2	S76043	dihydroxyacetone k
462	21	87.5	270	2	B69996	amino acid ABC tra	535	21	87.5	315	2	P86655	UDP-N-acetylmutama
463	21	87.5	270	2	F96971	CBIK protein (chal	536	21	87.5	316	2	F98027	UDP-N-acetylenolp
464	21	87.5	270	2	T35365	hypothetical prote	537	21	87.5	316	2	G95161	probable NADH2 deh
465	21	87.5	271	2	S18730	aminoglycoside N3'	538	21	87.5	316	2	S57850	NADH2 dehydrogen
466	21	87.5	272	2	G71339	probable amino aci	539	21	87.5	317	2	S68157	
467	21	87.5	273	2	AB2049	hypothetical prote	540	21	87.5	317	2		

541	21	87.5	317	2	AG2729	agmatinase [import	614	21	87.5	354	2	G64475	GTP-binding protei
542	21	87.5	317	2	A97511	hypothetical prote	615	21	87.5	354	2	T52401	branched-chain ami
543	21	87.5	318	2	T52663	thiosulfate sulfur	616	21	87.5	356	2	D87490	NADH dehydrogenase
544	21	87.5	318	2	T45180	hypothetical prote	617	21	87.5	356	2	S66348	cysteine proteinase
545	21	87.5	319	2	E82809	export protein Xf0	618	21	87.5	356	2	T30426	cathepsin-like pro
546	21	87.5	319	2	S03833	hypothetical prote	619	21	87.5	356	2	D96537	hypothetical prote
547	21	87.5	320	2	H84092	phage-related prot	620	21	87.5	357	2	E47411	ADPribosylarginine
548	21	87.5	320	2	S07384	diaminopimelate de	621	21	87.5	358	2	D89823	hypothetical prote
549	21	87.5	322	2	S01499	NADH2 dehydrogenas	622	21	87.5	358	2	H75084	hypothetical prote
550	21	87.5	323	2	A34284	NADH2 dehydrogenas	623	21	87.5	358	2	H71058	hypothetical prote
551	21	87.5	323	2	S62736	cathepsin-like cys	624	21	87.5	359	2	S67691	probable membrane
552	21	87.5	323	2	JC5691	cysteine proteinase	625	21	87.5	359	2	AF3184	beta-lactamase [im
553	21	87.5	323	2	E83860	biotin operon tran	626	21	87.5	359	2	S40746	hypothetical prote
554	21	87.5	323	2	F69454	signal-transducing	627	21	87.5	360	2	S59598	cysteine proteinase
555	21	87.5	323	2	AH3187	nitrotriacetate	628	21	87.5	360	2	T32554	hypothetical prote
556	21	87.5	324	2	T24819	hypothetical prote	629	21	87.5	361	1	KHBH	aleurain (EC 3.4.2
557	21	87.5	324	2	F70609	hypothetical prote	630	21	87.5	361	2	T12613	adenosylmethionine
558	21	87.5	325	1	B40358	NADH2 dehydrogenas	631	21	87.5	362	1	KHR2OG	oryzain (EC 3.4.22
559	21	87.5	325	1	DNWU1	NADH2 dehydrogenas	632	21	87.5	362	1	ZPBOC1	ubiquinol-cytochro
560	21	87.5	325	1	S49576	NADH2 dehydrogenas	633	21	87.5	362	2	B83963	hypothetical prote
561	21	87.5	325	1	T11116	NADH2 dehydrogenas	634	21	87.5	363	2	F69878	conserved hypothet
562	21	87.5	325	2	A69595	transcription repr	635	21	87.5	363	2	B82090	acetyltransferase,
563	21	87.5	325	2	T24737	hypothetical prote	636	21	87.5	364	2	E89894	conserved hypothet
564	21	87.5	326	2	S70596	NADH2 dehydrogenas	637	21	87.5	364	2	AE3196	conserved hypothet
565	21	87.5	326	2	S59101	NADH2 dehydrogenas	638	21	87.5	366	2	T52655	thiosulfate sulfur
566	21	87.5	326	2	B73439	hypothetical prote	639	21	87.5	367	2	S66692	hypothetical prote
567	21	87.5	326	2	E87042	conserved hypothet	640	21	87.5	368	2	F82570	DnaJ protein XF233
568	21	87.5	326	2	S55115	GLO1 protein - Yea	641	21	87.5	368	2	AC0647	tRNA (5-methylamin
569	21	87.5	327	2	E81686	transaldolase TC05	642	21	87.5	368	2	T15492	hypothetical prote
570	21	87.5	327	2	G15531	probable transaldo	643	21	87.5	369	2	S53843	NADH2 dehydrogenas
571	21	87.5	327	2	G72411	hypothetical prote	644	21	87.5	369	2	AF0162	putrescine-binding
572	21	87.5	327	2	G90139	deacetylase, proba	645	21	87.5	369	2	E84542	hypothetical prote
573	21	87.5	327	2	A83848	ferric ion ABC tra	646	21	87.5	370	2	T37282	probable cathelain
574	21	87.5	328	1	S25993	NADH2 dehydrogenas	647	21	87.5	372	2	C82237	tRNA (5-methylamin
575	21	87.5	328	2	T11227	NADH2 dehydrogenas	648	21	87.5	372	2	T00243	sopA protein - Bsc
576	21	87.5	328	2	T11958	NADH2 dehydrogenas	649	21	87.5	373	2	E70458	translational releas
577	21	87.5	328	2	S74645	blivverdin reducta	650	21	87.5	373	2	D90032	hypothetical prote
578	21	87.5	329	2	A86300	hypothetical prote	651	21	87.5	373	2	S65583	matting type protei
579	21	87.5	330	1	B42941	site-specific DNA-	652	21	87.5	374	2	G70947	hypothetical prote
580	21	87.5	330	2	S58931	NADH2 dehydrogenas	653	21	87.5	374	2	AH1825	protoporphyrin IX
581	21	87.5	331	1	DN0BU1	NADH2 dehydrogenas	654	21	87.5	375	2	C83317	tRNA methyltransfe
582	21	87.5	331	2	C72239	hypothetical prote	655	21	87.5	376	2	B84360	citrate synthase [
583	21	87.5	331	2	E82389	probable outer mem	656	21	87.5	377	2	T47471	cysteine proteinase
584	21	87.5	332	2	I40792	hypothetical prote	657	21	87.5	378	2	T30965	hypothetical prote
585	21	87.5	332	2	A45710	BRRF2 gene homolog	658	21	87.5	378	2	D90245	n2,N2 dimethylguan
586	21	87.5	332	2	D83519	probable transcrip	659	21	87.5	379	2	C81880	probable homoserin
587	21	87.5	333	2	S78136	NADH2 dehydrogenas	660	21	87.5	379	2	T01034	thiosulfate sulfur
588	21	87.5	334	2	T11890	NADH2 dehydrogenas	661	21	87.5	379	2	E71296	probable spermidin
589	21	87.5	336	2	E70413	NADH2 dehydrogenas	662	21	87.5	380	2	G85657	hypothetical prote
590	21	87.5	336	2	E98083	hypothetical prote	663	21	87.5	380	2	E90797	hypothetical prote
591	21	87.5	336	2	S69524	hypothetical prote	664	21	87.5	380	2	G69277	branched-chain ami
592	21	87.5	336	2	T04024	hypothetical prote	665	21	87.5	382	2	S51962	FUN49 protein - Ye
593	21	87.5	337	2	A81920	GTP-binding protei	666	21	87.5	382	2	F81298	probable decarboxy
594	21	87.5	338	2	A72544	probable agmatinas	667	21	87.5	383	2	B64858	Probable Atpase yc
595	21	87.5	338	2	F87267	PDZ domain family	668	21	87.5	383	2	E90829	hypothetical prote
596	21	87.5	338	2	C96775	hypothetical prote	669	21	87.5	383	2	C85687	hypothetical prote
597	21	87.5	339	2	F97853	NADH2 dehydrogenas	670	21	87.5	384	2	T00625	branched-chain ami
598	21	87.5	339	2	F71640	NADH2 dehydrogenas	671	21	87.5	387	2	G64489	hypothetical prote
599	21	87.5	339	2	B81162	signal peptidase I	672	21	87.5	387	2	F95059	transmembrane prot
600	21	87.5	340	2	G72596	hypothetical prote	673	21	87.5	387	2	H82176	carboxymorspermid
601	21	87.5	341	2	A11429	transcription regu	674	21	87.5	387	2	T30241	hypothetical prote
602	21	87.5	341	2	AG1803	transcription regu	675	21	87.5	387	2	G71097	probable amidohydr
603	21	87.5	343	2	D86198	cysteine proteinase	676	21	87.5	388	1	BVECAF	sopA protein - Bsc
604	21	87.5	343	2	AG0278	conserved hypothet	677	21	87.5	388	2	D91240	probable amino aci
605	21	87.5	344	2	S62765	NADH2 dehydrogenas	678	21	87.5	388	2	A86088	probable hippurica
606	21	87.5	344	2	T41659	probable potassium	679	21	87.5	390	2	B81408	probable periplasm
607	21	87.5	345	2	E87209	probable enoyl-CoA	680	21	87.5	391	1	D64801	probable monooxyge
608	21	87.5	346	2	A86866	hypothetical prote	681	21	87.5	391	2	AD0584	probable monooxyge
609	21	87.5	348	2	F83993	C4-dicarboxylate't	682	21	87.5	391	2	G85566	hypothetical prote
610	21	87.5	349	2	T45431	glycerol-3-phospha	683	21	87.5	391	2	D90716	oxygenase [importe
611	21	87.5	350	2	D69143	LPS biosynthesis R	684	21	87.5	391	2	D64366	hypothetical prote
612	21	87.5	351	2	F87119	glycerol-3-phospha	685	21	87.5	391	2	T28632	Y4cE protein - Rhi
613	21	87.5	352	2	F89884	Phe-tRNA synthetas	686	21	87.5	391	2	T40169	hypothetical prote

687	21	87.5	391	2	C96031	probable ureashort	760	21	87.5	445	2	D71977	phosphoglucosamine
688	21	87.5	393	2	S16844	titin - rabbit (fr	761	21	87.5	445	2	C38537	ureC protein - Hel
689	21	87.5	394	2	G82992	ubiH protein PA522	762	21	87.5	446	2	T45525	WSC4 homolog [impo
690	21	87.5	394	2	E72553	probable citrate s	763	21	87.5	448	2	B75180	antibiotic/peptide
691	21	87.5	394	2	A71083	hypothetical prote	764	21	87.5	448	2	T39314	hypothetical prote
692	21	87.5	397	2	T06531	pyruvate dehydroge	765	21	87.5	449	2	A54760	translation elonga
693	21	87.5	397	2	F82257	ViaC-related prote	766	21	87.5	449	2	JC5117	conserved hypothet
694	21	87.5	397	2	T11786	aspartate transami	767	21	87.5	450	2	D97885	glucosyl-1-phospha
695	21	87.5	397	2	AE1142	penicillin-binding	768	21	87.5	455	2	S52736	adenylosuccinate 1
696	21	87.5	398	2	A58922	esterase/N-deacety	769	21	87.5	456	2	A82237	S-adenosylmethioni
697	21	87.5	398	2	AH1500	penicillin-binding	770	21	87.5	456	2	H97076	mannitol 2-dehydro
698	21	87.5	400	2	AD2332	phosphoglycerate k	771	21	87.5	458	2	F87433	chromosome replica
699	21	87.5	402	2	E82071	ubiH protein VC247	772	21	87.5	458	2	AC1632	hypothetical prote
700	21	87.5	403	2	A82503	multidrug resistanc	773	21	87.5	458	2	T70325	hypothetical prote
701	21	87.5	404	2	H87635	Rieske 2Fe-2S fami	774	21	87.5	460	2	T21678	probable cytochrom
702	21	87.5	405	2	T25320	hypothetical prote	775	21	87.5	461	2	H70899	panthothenate perme
703	21	87.5	405	2	G55471	pyruvate decarboxy	776	21	87.5	461	2	E69279	hypothetical prote
704	21	87.5	405	2	G44450	isoleucine-tRNA li	777	21	87.5	461	2	D89894	probable glutamine
705	21	87.5	406	2	G84529	transposase TnpA h	778	21	87.5	462	2	A72521	transcription regu
706	21	87.5	406	2	A40629	amK protein - Erw	780	21	87.5	462	2	G83986	probable phosphogl
707	21	87.5	407	2	S52148	mRNA 3'-end proces	781	21	87.5	466	2	T41375	dep protein precur
708	21	87.5	407	2	B87262	2-octaprenyl-6-met	782	21	87.5	466	2	S36209	cysteinyI-tRNA syn
709	21	87.5	412	2	J00760	methyItransferase	783	21	87.5	466	2	C89820	acid phosphatase (
710	21	87.5	412	2	Q8HUB2	beta-2-adrenergic	784	21	87.5	467	1	PABYC	probable liposamide
711	21	87.5	413	2	T43170	probable triacylgly	785	21	87.5	467	2	A72548	triglyceride lipas
712	21	87.5	413	2	G84462	probable SCARECROW	786	21	87.5	467	2	T41053	acid phosphatase (
713	21	87.5	413	2	T35502	probable two-compo	787	21	87.5	467	2	S48996	acid phosphatase (
714	21	87.5	414	2	A75505	probable glutaryl-	788	21	87.5	467	2	S53476	3-isopropylmalate
715	21	87.5	414	2	AG2407	site-specific DNA-	789	21	87.5	467	2	H82070	drug resistance tr
716	21	87.5	415	2	T38324	probable trna meth	790	21	87.5	468	2	C81083	hypothetical prote
717	21	87.5	415	2	S68831	CWP-N-acetylneuram	791	21	87.5	468	2	E69294	probable phosphogl
718	21	87.5	415	2	S68831	hypothetical prote	792	21	87.5	469	2	S48370	probable phosphogl
719	21	87.5	416	2	AH2387	probable Affrase HI	793	21	87.5	469	2	B83722	transcription regu
720	21	87.5	418	1	D64144	hypothetical prote	794	21	87.5	469	2	G81128	3-isopropylmalate
721	21	87.5	420	2	H84360	I7 protein - fowlp	795	21	87.5	470	2	A81836	probable 3-isoprop
722	21	87.5	421	1	F48563	hypothetical prote	796	21	87.5	470	2	A81270	3-isopropylmalate
723	21	87.5	421	1	E71363	hypothetical prote	797	21	87.5	470	2	AD2435	hypothetical prote
724	21	87.5	423	1	WZVZ17	I7 protein - vacci	798	21	87.5	471	2	S11899	glutamate-ammonia
725	21	87.5	423	2	T11214	3-oxoacyl-lacyl-ca	799	21	87.5	471	2	A99180	hypothetical prote
726	21	87.5	423	2	H81200	cell division prot	800	21	87.5	471	2	S11899	hypothetical prote
727	21	87.5	423	2	T37344	topoisomerase II -	801	21	87.5	472	2	C71439	hypothetical prote
728	21	87.5	423	2	C42511	I7L protein - vacc	802	21	87.5	472	2	A70432	biotin carboxylase
729	21	87.5	423	2	E36843	K7L protein - vari	803	21	87.5	473	1	T40391	hypothetical prote
730	21	87.5	423	2	T28499	hypothetical prote	804	21	87.5	473	1	S75568	phosphoprotein pho
731	21	87.5	423	2	C72158	L7L protein - vari	805	21	87.5	474	1	S75568	isochorismate synt
732	21	87.5	424	1	S07792	site-specific DNA-	806	21	87.5	474	1	JC6042	cytosol nonspecific
733	21	87.5	425	2	H79964	pyrimidine-nucleos	807	21	87.5	474	1	T39587	serine/threonine-p
734	21	87.5	425	2	D95097	pyrimidine-nucleos	808	21	87.5	475	2	H82829	two-component syst
735	21	87.5	425	2	D97937	hypothetical prote	809	21	87.5	475	2	AG0801	probable transcrip
736	21	87.5	426	2	C90425	metabolite transpo	810	21	87.5	476	2	F70791	cytochrome P450 Rv
737	21	87.5	427	2	B90371	hypothetical prote	811	21	87.5	477	2	G70427	biotin carboxylase
738	21	87.5	429	2	S46801	TOM34 protein - Ye	812	21	87.5	479	2	T03293	probable phosphodi
739	21	87.5	429	2	S45459	folypolyglutamate	813	21	87.5	479	2	H90042	hypothetical prote
740	21	87.5	430	2	D72411	flagellin protein	814	21	87.5	480	2	C69438	hypothetical prote
741	21	87.5	430	2	AB2646	histidinol dehydro	815	21	87.5	482	2	E97748	virB10 protein [im
742	21	87.5	431	2	F82585	site-specific DNA-	816	21	87.5	483	2	F71684	virB10 protein (vi
743	21	87.5	431	2	AE3090	F14I17.11 protein	817	21	87.5	483	2	E71133	hypothetical prote
744	21	87.5	431	2	G86277	cell division prot	818	21	87.5	484	2	AI2527	hypothetical prote
745	21	87.5	435	2	C81777	probable 2-ketoglu	819	21	87.5	484	2	E84765	hypothetical prote
746	21	87.5	435	2	E83362	flad protein (U951	820	21	87.5	485	2	H72562	probable aldehyde
747	21	87.5	436	2	A97428	phosphate regulon	821	21	87.5	486	2	B91289	mannonate oxidore
748	21	87.5	438	2	AD0389	hypothetical prote	822	21	87.5	486	2	S56548	D-mannonate oxid
749	21	87.5	438	2	B91285	hypothetical prote	823	21	87.5	486	2	H90220	hypothetical prote
750	21	87.5	438	2	D86126	hypothetical prote	824	21	87.5	488	2	T13385	hypothetical prote
751	21	87.5	438	2	G87675	arylesterase-relat	825	21	87.5	488	2	A47259	corticosteroid-bin
752	21	87.5	441	2	D96802	hypothetical prote	826	21	87.5	489	2	T13385	hypothetical prote
753	21	87.5	443	2	T39540	triglyceride lipas	827	21	87.5	490	2	AH0883	D-mannonate oxid
754	21	87.5	443	2	F90213	hypothetical prote	828	21	87.5	490	2	A46391	cAMP receptor subt
755	21	87.5	443	2	T04025	hypothetical prote	829	21	87.5	490	2	D87644	conserved hypothet
756	21	87.5	443	2	A54813	cAMP receptor CAR4	830	21	87.5	490	2	F87443	conserved hypothet
757	21	87.5	444	2	A70674	probable mbtC prot	831	21	87.5	491	2	T19505	hypothetical prote
758	21	87.5	444	2	F97750	colB protein precu	832	21	87.5	495	2	B70322	hypothetical prote
759	21	87.5	445	1	F71033	hypothetical prote							

833	21	87.5	495	2	E69731	PBSX prophage ORF	906	21	87.5	564	2	S74503	hypothetical prote
834	21	87.5	496	2	I39478	repr protein - Str	907	21	87.5	564	2	S50700	pyruvate decarboxy
835	21	87.5	496	2	S06925	replication initia	908	21	87.5	566	1	F64329	arginine-trRNA liga
836	21	87.5	496	2	JH0206	hypothetical 57.4K	909	21	87.5	567	2	D84400	phenylalanyl-tRNA
837	21	87.5	497	1	B42902	methylmalonate-sem	910	21	87.5	568	1	DCZY22	pyruvate decarboxy
838	21	87.5	497	2	JC4924	aldehyde dehydroge	911	21	87.5	569	2	B87180	pyruvate (or indol
839	21	87.5	498	2	G37567	monooxygenase, fla	912	21	87.5	570	1	JN0782	pyruvate decarboxy
840	21	87.5	499	2	S43324	zeta-carotene desa	913	21	87.5	570	2	T11647	probable pyruvate
841	21	87.5	499	2	AG2509	zeta-carotene desa	914	21	87.5	571	2	T38759	probable pyruvate
842	21	87.5	500	2	T51172	betaine-aldehyde d	915	21	87.5	572	2	T51575	2-hydroxyphytanoyl
843	21	87.5	500	2	B72522	hypothetical prote	916	21	87.5	572	2	T50136	probable pyruvate
844	21	87.5	501	2	A59090	aspartic proteinas	917	21	87.5	573	2	H89904	aerobic glycerol-3
845	21	87.5	501	2	T32642	hypothetical prote	918	21	87.5	573	2	F86663	hypothetical prote
846	21	87.5	502	2	G70193	hypothetical prote	919	21	87.5	574	2	T25887	hypothetical prote
847	21	87.5	503	2	T39216	aldehyde dehydroge	920	21	87.5	574	2	G70468	DNA polymerase I (
848	21	87.5	503	2	T13006	betaine aldehyde d	921	21	87.5	575	2	S57648	hypothetical ABC t
849	21	87.5	504	1	G71248	tldd homolog PH024	922	21	87.5	575	2	G98350	hypothetical ABC t
850	21	87.5	506	2	B87693	aldehyde dehydroge	923	21	87.5	576	2	T12896	probable single-st
851	21	87.5	510	2	B98381	protein T22P7.1 [i	924	21	87.5	577	2	A57469	CMP-N-acetylneuram
852	21	87.5	511	2	G36688	hypothetical prote	925	21	87.5	585	2	T03252	pyruvate decarboxy
853	21	87.5	513	2	C81859	probable integral	926	21	87.5	585	2	G96995	ATP-dependent RNA
854	21	87.5	514	2	C90060	1-pyrroline-5-carb	927	21	87.5	587	2	D84426	hypothetical prote
855	21	87.5	520	2	B33830	cation efflux syst	928	21	87.5	588	2	AI3543	60K inner membrane
856	21	87.5	520	2	JC4699	cadmium, zinc, cob	929	21	87.5	590	2	S72961	CTP synthase (EC 6
857	21	87.5	522	2	S06431	2-isopropylmalate	930	21	87.5	591	2	E81687	ATP synthase, chai
858	21	87.5	522	2	S71821	probable interleuk	931	21	87.5	591	2	B71531	probable ATP synth
859	21	87.5	523	2	AF0516	2-isopropylmalate	932	21	87.5	592	2	T48155	pyruvate decarboxy
860	21	87.5	523	2	S11309	glucuronosyltransf	933	21	87.5	593	2	S65470	pyruvate decarboxy
861	21	87.5	523	2	S38186	hypothetical prote	934	21	87.5	594	2	T06369	probable 3,4-dihyd
862	21	87.5	523	2	T51475	RGA-like protein -	935	21	87.5	594	2	T38114	pyruvate decarboxy
863	21	87.5	527	2	T21830	hypothetical prote	936	21	87.5	598	2	S35255	2,4-dichlorophenol
864	21	87.5	528	2	C83033	benzoylformate dec	937	21	87.5	598	2	H64496	hypothetical prote
865	21	87.5	528	2	JN0619	glucuronosyltransf	938	21	87.5	599	2	T10798	phosphorin-S - Vo
866	21	87.5	528	2	D97270	ATP-dependent RNA	939	21	87.5	602	2	S71557	pyruvate decarboxy
867	21	87.5	528	2	C87663	conserved hypothet	940	21	87.5	603	2	T03295	pyruvate decarboxy
868	21	87.5	529	2	J50200	orphan UDP-glucuro	941	21	87.5	603	2	T48154	pyruvate decarboxy
869	21	87.5	530	2	E82491	sigma-54 dependent	942	21	87.5	604	2	E75119	hypothetical prote
870	21	87.5	530	2	S68200	glucuronosyltransf	943	21	87.5	605	2	H69581	transcription acti
871	21	87.5	531	2	S20900	titin - mouse (fra	944	21	87.5	605	2	T43191	probable pyruvate
872	21	87.5	531	2	B84442	hypothetical prote	945	21	87.5	606	2	T27072	hypothetical prote
873	21	87.5	532	2	T34917	peptide synthase -	946	21	87.5	607	2	T05315	pyruvate decarboxy
874	21	87.5	533	2	H86282	protein F1086.34 [	947	21	87.5	608	2	H71379	probable tpr prote
875	21	87.5	533	2	E84858	phosphoprotein pho	948	21	87.5	608	2	A64992	sensor protein Ato
876	21	87.5	537	2	E75009	hypothetical prote	949	21	87.5	609	2	AD0917	ATP-dependent DNA
877	21	87.5	537	2	A46611	myosin-binding pro	950	21	87.5	609	2	F84824	hypothetical prote
878	21	87.5	539	2	D71260	hypothetical prote	951	21	87.5	609	2	G71731	mitochondrial tran
879	21	87.5	540	2	JG0174	phospholipase D (E	952	21	87.5	610	1	BVECRQ	DNA helicase recQ
880	21	87.5	542	2	S45557	resB protein - Bac	953	21	87.5	610	1	DCZMP	pyruvate decarboxy
881	21	87.5	545	2	T01288	protein kinase F27	954	21	87.5	610	1	AL0466	ATP-dependent DNA
882	21	87.5	546	2	F95038	hypothetical prote	955	21	87.5	611	2	F86069	ATP-dependent DNA
883	21	87.5	547	2	AD2250	acetylactate synth	956	21	87.5	611	2	H91222	ATP-dependent DNA
884	21	87.5	547	2	S70538	signal-transducing	957	21	87.5	611	2	D81347	arginine decarboxy
885	21	87.5	548	2	AG2931	hypothetical prote	958	21	87.5	611	2	E90543	glucose inhibited
886	21	87.5	550	2	S75122	acetylactate synth	959	21	87.5	611	2	F70325	conserved hypothet
887	21	87.5	550	2	B82912	arginyl tRNA synth	960	21	87.5	612	2	L64241	glucose inhibited
888	21	87.5	550	2	AG0095	probable sideropho	961	21	87.5	612	2	S73611	glucose inhibited
889	21	87.5	552	2	A71864	ribosomal protein	962	21	87.5	612	2	T38714	hypothetical prote
890	21	87.5	552	2	F96769	hypothetical prote	963	21	87.5	614	2	D87410	old protein [impo
891	21	87.5	552	2	T16345	hypothetical prote	964	21	87.5	614	2	S57821	pyruvate decarboxy
892	21	87.5	555	2	C45868	glycerol-3-phospha	965	21	87.5	617	2	S74002	translation initia
893	21	87.5	556	2	G64569	ribosomal protein	966	21	87.5	617	2	F84681	hypothetical prote
894	21	87.5	558	2	AE1236	glycerol 3 phospho	967	21	87.5	619	2	G64215	translation initia
895	21	87.5	558	2	AB1599	glycerol 3 phospho	968	21	87.5	619	2	F81324	glucose inhibited
896	21	87.5	558	2	T34849	probable acetyl-co	969	21	87.5	620	2	B82351	ATP-dependent DNA
897	21	87.5	559	1	DCZYPC	pyruvate decarboxy	970	21	87.5	621	2	E64546	glucose inhibited
898	21	87.5	560	1	F69059	arginine-trRNA liga	971	21	87.5	621	2	D71961	glucose inhibited
899	21	87.5	562	2	T49386	hypothetical prote	972	21	87.5	621	2	H71713	glucose inhibited
900	21	87.5	562	2	S46281	p element - fruit	973	21	87.5	622	2	D97710	glucose inhibited
901	21	87.5	563	1	DCBYP	pyruvate decarboxy	974	21	87.5	622	2	PT0059	phosphinothricin-t
902	21	87.5	563	1	S70684	pyruvate decarboxy	975	21	87.5	624	2	D97307	molecular chaperon
903	21	87.5	563	2	S64382	pyruvate decarboxy	976	21	87.5	625	2	H75110	arginyl-tRNA synth
904	21	87.5	563	2	S59324	pyruvate decarboxy	977	21	87.5	625	2	F90080	glucose inhibited
905	21	87.5	564	1	S36363	pyruvate decarboxy	978	21	87.5	625	2	G86855	glucose inhibited

```
979      21 87.5 626 2 H69642
980      21 87.5 627 2 AF2924
981      21 87.5 627 2 E97698
982      21 87.5 627 2 T27123
983      21 87.5 628 1 BWSGA
984      21 87.5 628 2 A84930
985      21 87.5 628 2 A81939
986      21 87.5 629 1 BVECOA
987      21 87.5 629 2 AF0953
988      21 87.5 629 2 C82598
989      21 87.5 629 2 C91214
990      21 87.5 629 2 D86060
991      21 87.5 629 2 I64078
992      21 87.5 629 2 AF0501
993      21 87.5 629 2 A11425
994      21 87.5 629 2 AG1799
995      21 87.5 629 2 E69284
996      21 87.5 629 2 A71023
997      21 87.5 630 1 BWPSAP
998      21 87.5 630 2 D82950
999      21 87.5 631 2 F81227
1000     21 87.5 631 2 G82035

                                ALIGNMENTS

RESULT 1
JU0240
nuclear matrix protein N/MAX-74 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C:Accession: JU0240
R:Inagaki, H.; Matsushima, Y.; Ohshima, M.; Nakamura, K.; Kadowaki, T.; Kitagawa, Y.
submitted to JIPID, October 1993
A:Description: Twice repeated amino acid sequence segments common to Matrin 3 and a novel
A:Reference number: JU0240
A:Accession: JU0240
A:Molecule type: protein
A:Residues: 1-74 <INA>
A:Cross-references: UNIPROT:Q14966
A:Experimental source: Namalwa cell; Hela cell
C:Keywords: nuclear matrix

Query Match          91.7%; Score 22; DB 2; Length 74;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      |||
Db      18 GYSVE 22

RESULT 2
AI3049
transcription regulator Atu4014 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AI3049
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AI3049
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-76 <KUR>
A:Cross-references: UNIPROT:QBU856; GB:AE008689; PIDN:AAL44815.1; PID:gl7742457; GSPDB:G
A:Experimental source: strain C58 (Dupont)
```

```
C:Genetics:
A:Gene: Atu4014
A:Map position: linear chromosome

Query Match          91.7%; Score 22; DB 2; Length 76;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      |||
Db      30 GYVE 34

RESULT 3
RDECD5
dihydrofolate reductase (EC 1.5.1.3) type II - Escherichia coli plasmid R751
N:Alternate names: tetrahydrofolate dehydrogenase
C:Species: Escherichia coli
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C:Accession: A23598
R:Plensburg, J.; Steen, R.
Nucleic Acids Res. 14, 5933, 1986
A:Title: Nucleotide sequence analysis of the trimethoprim resistant dihydrofolate reductase
A:Reference number: A23598; MUID:86286598; PMID:3526286
A:Accession: A23598
A:Molecule type: DNA
A:Residues: 1-78 <FLE>
A:Cross-references: UNIPROT:P05794; GB:X04128; NID:g41268; PIDN:CAA27740.1; PID:g41269
A:Note: the authors translated the codon TGC for residue 47 as Lys
C:Genetics:
A:Genome: plasmid
C:Superfamily: dihydrofolate reductase, type II
C:Keywords: NADP; oxidoreductase; trimethoprim resistance

Query Match          91.7%; Score 22; DB 1; Length 78;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      |||
Db      54 GYAVE 58

RESULT 4
RDECD6
dihydrofolate reductase (EC 1.5.1.3) type II - Escherichia coli plasmid R67
C:Species: Escherichia coli
C>Date: 31-Mar-1980 #sequence_revision 29-Jul-1981 #text_change 09-Jul-2004
C:Accession: A91512; A92244; A00397; A24948
R:Brisson, N.; Hohn, T.
Gene 28, 271-274, 1984
A:Title: Nucleotide sequence of the dihydrofolate-reductase gene borne by the plasmid R67
A:Reference number: A91512; MUID:84237581; PMID:6735180
A:Accession: A91512
A:Molecule type: DNA
A:Residues: 1-78 <BRI>
A:Cross-references: UNIPROT:P00383; GB:X02118; GB:M27532; NID:gl151839; PIDN:AAA26083.1; I
R:Stone, D.; Smith, S.L.
J. Biol. Chem. 254, 10857-10861, 1979
A:Title: The amino acid sequence of the trimethoprim-resistant dihydrofolate reductase of
A:Reference number: A92244; MUID:80049683; PMID:387758
A:Accession: A92244
A:Molecule type: protein
A:Residues: 1-78 <STO>
A:Note: this protein is specified by plasmid R67 expressed in Escherichia coli
C:Comment: Type II plasmid-specified enzyme is practically insensitive to trimethoprim ar
C:Genetics:
A:Genome: plasmid
C:Superfamily: dihydrofolate reductase, type II
C:Keywords: NADP; oxidoreductase

Query Match          91.7%; Score 22; DB 1; Length 78;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
```



Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 54 GYAVE 58

RESULT 5  
RDECD8  
C:Species: Escherichia coli  
C>Date: 29-Jul-1981 #sequence\_revision 29-Jul-1981 #text\_change 09-Jul-2004  
C:Accession: A00398  
R:Zolg, J.W.; Hanggi, U.J.  
Nucleic Acids Res. 9, 697-710, 1981  
A:Title: Characterization of a R plasmid-associated, trimethoprim-resistant dihydrofolate  
A:Reference number: A00398; MUID:81174731; PMID:6261228  
A:Accession: A00398  
A:Molecule type: DNA  
A:Residues: 1-78 <ZOL>  
A:Cross-references: UNIPROT:P00384; GB:V00252; GB:J01774; NID:G40870; PIDN:CAA23503.1; F  
A:Note: this protein is specified by plasmid R388 expressed in Escherichia coli  
C:Comment: Type II plasmid-specified enzyme is practically insensitive to trimethoprim a  
C:Genetics:  
A:Genome: plasmid  
C:Superfamily: dihydrofolate reductase, type II  
C:Keywords: NADP; oxidoreductase

Query Match 91.7%; Score 22; DB 1; Length 78;  
Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 54 GYAVE 58

RESULT 6  
T08517  
C:Species: Enterobacter aerogenes  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 27-Oct-2003  
C:Accession: T08517  
R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.  
Plasmid 36, 95-111, 1996  
A:Title: Conservation of the genetic switch between replication and transfer genes of In  
A:Reference number: Z16434; MUID:97118926; PMID:8954881  
A:Accession: T08517  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-78 <THO>  
A:Cross-references: EMBL:U67194; NID:G1572520; PIDN:AAC64461.1; PID:G1572566  
C:Genetics:  
A:Gene: dhfrIIC  
A:Genome: plasmid R751  
C:Superfamily: dihydrofolate reductase, type II  
C:Keywords: NADP; oxidoreductase

Query Match 91.7%; Score 22; DB 2; Length 78;  
Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 54 GYAVE 58

RESULT 7  
S32183  
C:Species: Klebsiella pneumoniae  
C>Date: 03-Mar-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S32183

R:Radstroem, P.; Sundstrom, L.; Swedberg, G.; Flensburg, J.; Skold, O.  
submitted to the EMBL Data Library, March 1993  
A:Description: The integrons of R751 and Tn21 are transposons related to the retroelement  
A:Reference number: S32177  
A:Accession: S32183  
A:Molecule type: DNA  
A:Residues: 1-78 <RAD>  
A:Cross-references: UNIPROT:P05794; EMBL:X72585; NID:G288626; PIDN:CAAS1181.1; PID:G28862  
C:Superfamily: dihydrofolate reductase, type II  
C:Keywords: NADP; oxidoreductase

Query Match 91.7%; Score 22; DB 2; Length 78;  
Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 54 GYAVE 58

RESULT 8  
S58349  
N:Alternating two-sector ATPase (EC 3.6.3.14) chain H - Chlamydomonas reinhardtii chlor  
H\*-transporting two-sector ATPase (EC 3.6.3.14) chain H  
C:Species: chloroplast Chlamydomonas reinhardtii  
C>Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 09-Jul-2004  
C:Accession: S58349; S68396  
R:Rolland, N.; Rochaix, J.D.  
submitted to the EMBL Data Library, August 1995  
A:Reference number: S58348  
A:Accession: S58349  
A:Molecule type: DNA  
A:Residues: 1-82 <ROL>  
A:Cross-references: UNIPROT:Q37304; EMBL:X90559; NID:G944995; PIDN:CAA62149.1; PID:G94499  
A:Experimental source: wild type  
R:Fiedler, H.R.; Schmid, R.; Leu, S.; Shavit, N.; Strotmann, H.  
FEBS Lett. 377, 163-166, 1995  
A:Title: Isolation of CP(0)CF(1) from Chlamydomonas reinhardtii cw15 and the N-terminal  
A:Reference number: S68388; MUID:96128220; PMID:8543042  
A:Accession: S68396  
A:Molecule type: protein  
A:Residues: 1-7, 'A', 9-32 <FIE>  
A:Experimental source: strain CW15  
C:Genetics:  
A:Gene: atpH  
A:Genome: chloroplast  
C:Superfamily: H\*-transporting ATP synthase lipid-binding protein  
C:Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thylak  
F:7-33/Domain: transmembrane #status predicted <TM1>  
F:51-77/Domain: transmembrane #status predicted <TM2>  
F:1/Modified site: N-formylmethionine #status predicted  
F:61/Active site: Glu #status predicted

Query Match 91.7%; Score 22; DB 2; Length 82;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 33 GYAVE 37

RESULT 9  
T07211  
H\*-transporting two-sector ATPase (EC 3.6.3.14) lipid-binding protein - Chlorella vulgaris  
C:Species: chloroplast Chlorella vulgaris  
C>Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
C:Accession: T07211  
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakas  
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo  
A:Reference number: Z15985; MUID:97303241; PMID:9159184  
A:Accession: T07211

```

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-82 <WAK>
A;Cross-references: UNIPROT:P56297; EMBL:AB001684; NID:g2224352; PIDN:BAA57858.1; PID:g2224352
C;Genetics:
A;Gene: atph
A;Genome: chloroplast
A;Superfamily: H+-transporting ATP synthase lipid-binding protein
C;Keywords: chloroplast; hydrolase

Query Match          91.7%; Score 22; DB 2; Length 82;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
    ||||
Db 33 GYAVE 37

RESULT 10
C97083
uncharacterized protein, ortholog of YKUS B. subtilis [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
A;Accession: C97083
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: C97083
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-86 <KUR>
A;Cross-references: UNIPROT:Q97J00; GB:AE001437; PIDN:AAK79454.1; PID:gl5024432; GSPDB:G97J00
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1486

Query Match          91.7%; Score 22; DB 2; Length 86;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
    ||||
Db 21 GYSVE 25

RESULT 11
C37202
hypothetical protein (XynC 3' region) - Caldocellum saccharolyticum
C;Species: Caldocellum saccharolyticum
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 09-Jul-2004
A;Accession: C37202
R;Luethi, E.; Love, D.R.; McAnulty, J.; Wallace, C.; Caughey, P.A.; Saul, D.; Bergquist, A.; Appl. Environ. Microbiol. 56, 1017-1024, 1990
A;Title: Cloning, sequence analysis, and expression of genes encoding xylan-degrading enzymes from Caldocellum saccharolyticum
A;Reference number: A37202; MUID:90253140; PMID:2111111
A;Accession: C37202
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-97 <LUB>
A;Cross-references: UNIPROT:P23554; GB:M34459; NID:gl44295; PIDN:AAA23061.1; PID:gl44298

Query Match          91.7%; Score 22; DB 2; Length 97;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
    ||||
Db 69 GYSVE 73

RESULT 12
S64602
hypothetical protein YGR269w - Yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G9355
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 09-Jul-2004
A;Accession: S64602; S47510
R;Panzeri, L.; Agostoni Carbone, M.L.; Melchiorretto, P.; Plevani, P.; Martegani, E.; Vanc submitted to the Protein Sequence Database, May 1996
A;Reference number: S64591
A;Accession: S64602
A;Molecule type: DNA
A;Residues: 1-108 <PAN>
A;Cross-references: UNIPROT:P40326; EMBL:Z73053; NID:gl323487; PIDN:CAA97299.1; PID:gl323487
A;Experimental source: strain S288C
R;Mannhaupt, G.; Schnall, R.; Stucka, R.; Schwarze, C.; Vetter, I.; Feldmann, H. submitted to the EMBL Data Library, July 1994
A;Reference number: S47509
A;Accession: S47510
A;Molecule type: DNA
A;Residues: 1-52, 'A', '54-108 <MAN>
A;Cross-references: EMBL:X81072; NID:gl272679; PIDN:CAA56962.1; PID:g531765
C;Genetics:
A;Cross-references: SGD:S0003501
A;Map position: 7R
C;Superfamily: Saccharomyces hypothetical protein YGR269w

Query Match          91.7%; Score 22; DB 2; Length 108;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
    ||||
Db 44 GYTVE 48

RESULT 13
G71077
hypothetical protein PH0887 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
A;Accession: G71077
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon Pyrococcus horikoshii
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: G71077
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-109 <KAW>
A;Cross-references: UNIPROT:O58646; GB:AP000004; NID:g3236131; PIDN:BAA29981.1; PID:d1013
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0887

Query Match          91.7%; Score 22; DB 2; Length 109;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
    ||||
Db 21 GYTVE 25

RESULT 14
B98236
exs1 protein protein (AJ225561) [imported] - Agrobacterium tumefaciens (strain CS8, Cercariae)
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
A;Accession: B98236
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
```

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: B98236  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-119 <KUR>  
A;Cross-references: UNIPROT:Q8U8S6; GB:AE007870; PIDN:AAK89412.1; PID:g15159269; GSPDB:G000000000  
C;Genetics:  
A;Gene: AGR\_L1677  
A;Map position: linear chromosome

Query Match 91.7%; Score 22; DB 2; Length 119;  
Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 73 GYVE 77

RESULT 15  
A69453  
hypotheical protein AF1626 - *Archaeoglobus fulgidus*  
C;Species: *Archaeoglobus fulgidus*  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon *Archaeoglobus fulgidus*  
A;Reference number: A69250; MUID:98049343; PMID:9389475  
A;Accession: A69453  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-120 <KLE>  
A;Cross-references: UNIPROT:O28647; GB:AE000990; GB:AE000782; NID:g2689313; PIDN:AAB8962

Query Match 91.7%; Score 22; DB 2; Length 120;  
Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 19 GYAVE 23

RESULT 16  
B87323  
chemotaxis protein CheYIV [imported] - *Caulobacter crescentus*  
C;Species: *Caulobacter crescentus*  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 16-Aug-2004  
R;Nierman, W.C.; Reiblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of *Caulobacter crescentus*  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: B87323  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-122 <STO>  
A;Cross-references: UNIPROT:Q9AAK1; GB:AE005673; NID:g13421798; PIDN:AAK22582.1; GSPDB:G000000000  
C;Genetics:  
A;Gene: CC0596  
C;Superfamily: Signal transduction receiver (phosphoacceptor) protein, CheY type; response

Query Match 91.7%; Score 22; DB 2; Length 122;

Best Local Similarity 80.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 25 GYAVE 29

RESULT 17  
C96604  
hypotheical protein F14G9.i3 [imported] - *Arabidopsis thaliana*  
C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
R;Accession: C96604  
C;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: C96604  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-123 <STO>  
A;Cross-references: UNIPROT:Q9C7K1; GB:AE005173; NID:g11094714; PIDN:AAK29649.1; GSPDB:G000000000  
C;Genetics:  
A;Gene: F14G9.i3  
A;Map position: 1

Query Match 91.7%; Score 22; DB 2; Length 123;  
Best Local Similarity 80.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 38 GYSVE 42

RESULT 18  
D84401  
30S ribosomal protein S6E [imported] - *Halobacterium* sp. NRC-1  
C;Species: *Halobacterium* sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: D84401  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;  
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A;Title: Genome sequence of *Halobacterium* species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950  
A;Accession: D84401  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-131 <STO>  
A;Cross-references: UNIPROT:Q9HMJ5; GB:AE004437; NID:g10581909; PIDN:AAK20576.1; GSPDB:G000000000  
C;Genetics:  
A;Gene: rps6e  
C;Superfamily: Haloarcula ribosomal protein HS13

Query Match 91.7%; Score 22; DB 2; Length 131;  
Best Local Similarity 80.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 47 GYVE 51

RESULT 19  
H31194  
hypothetical protein PA3611 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: H31194  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizioch, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 953-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: H31194  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-136 <STO>  
A:Cross-references: UNIPROT:Q9HY15; GB:AE004781; GB:AE004091; NID:g9949761; PIDN:AA060699  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3611

Query Match 91.7%; Score 22; DB 2; Length 136;  
Best Local Similarity 80.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 54 GYTVE 58

RESULT 20  
D75265  
conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: D75265  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Shen, M.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: D75265  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-136 <WHI>  
A:Cross-references: UNIPROT:Q9RR12; GB:AE002080; GB:AE000513; NID:g6460327; PIDN:AAF1205  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR2509  
A:Map position: 1

Query Match 91.7%; Score 22; DB 2; Length 136;  
Best Local Similarity 80.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 98 GYTVE 102

RESULT 21  
G69473  
hypothetical protein AF1792 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: G69473  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: G69473  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-137 <KLE>  
A:Cross-references: UNIPROT:O28482; GB:AE000979; GB:AE000782; NID:g2689302; PIDN:AA089473

Query Match 91.7%; Score 22; DB 2; Length 137;  
Best Local Similarity 80.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 36 GYAVE 40

RESULT 22  
H72520  
hypothetical protein APE2141 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: H72520  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: H72520  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-142 <KAW>  
A:Cross-references: UNIPROT:Q9Y9Z8; DBJ:AP000063; NID:g5105654; PIDN:BAA81152.1; PID:g5105654  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE2141  
C:Superfamily: Aeropyrum pernix hypothetical protein APE2141

Query Match 91.7%; Score 22; DB 2; Length 142;  
Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 89 GYTVE 93

RESULT 23  
AF1081  
50S ribosomal protein L9 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AF1081  
R:Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AF1081  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-148 <GLA>  
A:Cross-references: UNIPROT:Q8VAR2; GB:NC\_003210; PIDN:CAC98268.1; PID:g16409412; GSPDB:C  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: rplI  
C:Superfamily: Escherichia coli ribosomal protein L9

Query Match 91.7%; Score 22; DB 2; Length 148;

Best Local Similarity 80.0%; Pred. No. 2.4e+02; Mismatches 1; Indels 0; Gaps 0;  
 Matches 4; Conservative 0;

Qy 1 GYXVE 5  
 ||||  
 Db 34 GYAVE 38

RESULT 24  
 AG1438  
 50S ribosomal protein L9 [imported] - *Listeria innocua* (strain Clip11262)  
 C:Species: *Listeria innocua*  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AG1438  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, W.; Zimmermann, P.; Zisch, A.  
 A:Title: Comparative genomics of *Listeria species*.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AG1438  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-148 <GLA>  
 A:Cross-references: UNIPROT:Q92FQ7; GB:AL592022; PIDN:CAC95279.1; PID:g16412467; GSPDB:Q92FQ7  
 A:Experimental source: strain Clip11262  
 C:Genetics:  
 A:Gene: rplI  
 C:Superfamily: Escherichia coli ribosomal protein L9

Query Match 91.7%; Score 22; DB 2; Length 148;  
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 34 GYAVE 38

RESULT 25  
 G75066  
 hypothetical protein PAB1006 - *Pyrococcus abyssi* (strain Orsay)  
 C:Species: *Pyrococcus abyssi*  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C:Accession: G75066  
 R:anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure  
 A:Reference number: A75001  
 A:Accession: G75066  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-148 <KAW>  
 A:Cross-references: UNIPROT:Q9UYI4; GB:AJ248287; GB:AL096836; NID:gs458657; PIDN:CAB5042  
 A:Experimental source: strain Orsay  
 C:Genetics:  
 A:Gene: PAB1006  
 C:Superfamily: Methanococcus jannaschii hypothetical protein MJ0648

Query Match 91.7%; Score 22; DB 2; Length 148;  
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 86 GYAVE 90

RESULT 26  
 B71166  
 hypothetical protein PH0527 - *Pyrococcus horikoshii*

C:Species: *Pyrococcus horikoshii*  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004  
 C:Accession: B71166  
 R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: B71166  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-151 <KAW>  
 A:Cross-references: GB:AP000002; NID:g3236129; PIDN:BA29615.1; PID:g3256932  
 A:Experimental source: strain Orf3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH0527

Query Match 91.7%; Score 22; DB 2; Length 151;  
 Best Local Similarity 80.0%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 89 GYAVE 93

RESULT 27  
 B64485  
 hypothetical protein MJ1483 - *Methanococcus jannaschii*  
 C:Species: *Methanococcus jannaschii*  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
 C:Accession: B64485  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Raon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C. A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*.  
 A:Reference number: A64300; MUID:96337999; PMID:8688087  
 A:Accession: B64485  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-152 <BUL>  
 A:Cross-references: UNIPROT:Q58878; GB:U67589; GB:L77117; NID:g2826417; PIDN:AAB99498.1;  
 C:Genetics:  
 A:Map position: FOR1454330-1454788  
 A:Start codon: TTG  
 C:Superfamily: Methanococcus jannaschii hypothetical protein MJ1483

Query Match 91.7%; Score 22; DB 2; Length 152;  
 Best Local Similarity 80.0%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 107 GYSVE 111

RESULT 28  
 S19721  
 pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) chain E1-beta - *Staphylococcus aureus* (f1)  
 C:Species: *Staphylococcus aureus*  
 C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 12-Jul-2004  
 C:Accession: S19721  
 R:Hemilae, H.  
 Biochim. Biophys. Acta 1129, 119-123, 1991  
 A:Title: Lipoamide dehydrogenase of *Staphylococcus aureus*: nucleotide sequence and sequencing  
 A:Reference number: S19721; MUID:92096451; PMID:1756171  
 A:Accession: S19721  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-154 <HEM>

```
A;Cross-references: UNIPROT:Q59820; EMBL:X59434; NID:G48871; PIDN:CAA41338.1; PID:G48872
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991
C;Superfamily: pyruvate dehydrogenase, E1 component, beta subunit
C;Keywords: oxidoreductase

Query Match          91.7%; Score 22; DB 2; Length 154;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 57 GYSVE 61

RESULT 29
PH0201
hypotheical protein 157 - Enterococcus faecalis plasmid pAM-beta-1 (fragment)
C;Species: Enterococcus faecalis
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: PH0201
R;Swinfieid, T.J.; Oultram, J.D.; Thompson, D.E.; Brehm, J.K.; Minton, N.P.
Gene 87, 79-90, 1990
A;Title: Physical Characterisation of the replication region of the Streptococcus faecalis
A;Reference number: PH0201; MUID:90236302; PMID:2110101
A;Accession: PH0201
A;Molecule type: DNA
A;Residues: 1-157 <SWI>
A;Cross-references: UNIPROT:Q52245; EMBL:X17092; NID:G3023041; PIDN:AAC38598.1; PID:G302
C;Genetics:
C;Keywords: plasmid replication

Query Match          91.7%; Score 22; DB 2; Length 157;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 112 GYSVE 116

RESULT 30
DB3384
probable transcription regulator PA2082 [imported] - Pseudomonas aeruginosa (strain PA01
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: DB3384
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: DB3384
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-158 <STO>
A;Cross-references: UNIPROT:Q91233; GB:AE004636; GB:AE004091; NID:G9948093; PIDN:AAG0547
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2082
C;Superfamily: regulatory protein asnC

Query Match          91.7%; Score 22; DB 2; Length 158;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 51 GYSVE 55

RESULT 31
```

```
H69504
hypotheical protein AF2041 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: H69504
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.A
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: H69504
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Cross-references: UNIPROT:O28238; GB:AE000962; GB:AE000782; NID:G2689285; PIDN:AAB89222
A;Residues: 1-162 <KLE>
A;Experimental source: strain A3(2)

Query Match          91.7%; Score 22; DB 2; Length 162;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 128 GYAVE 132

RESULT 32
T35779
probable secreted protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T35779
R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1998
A;Reference number: Z21570
A;Accession: T35779
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-168 <SEE>
A;Cross-references: UNIPROT:O87846; EMBL:AL031013; PIDN:CAA19784.1; GSPDB:GN00070; SCOREDI
C;Genetics:
C;Superfamily: Streptomyces coelicolor probable secreted protein SC8A6.12c

Query Match          91.7%; Score 22; DB 2; Length 168;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 47 GYAVE 51

RESULT 33
H70702
hypotheical protein RV2311 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: H70702
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70702
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-174 <COL>
```

```

A;Cross-references: UNIPROT:P71901; GB:Z79702; GB:AL123456; NID:g3261642; PIDN:CAB02056.
A;Experimental source: strain H37Rv
C;Genetics:
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv2311

Query Match          91.7%; Score 22; DB 2; Length 174;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
    ||||
Db 20 GYSVE 24

RESULT 34
A70395
hypothetical protein aq_1103 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: A70395
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; PMID:98196666; PMID:9537320
A;Accession: A70395
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-174 <AQF>
A;Cross-references: UNIPROT:Q67188; GB:AE000723; NID:g2983569; PIDN:AAC07152.1; PID:g298
A;Experimental source: strain VFS
C;Genetics:
A;Gene: aq_1103

Query Match          91.7%; Score 22; DB 2; Length 174;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
    ||||
Db 143 GYAVE 147

RESULT 35
B71274
probable translation initiation factor 3 (infC) - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: B71274
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; PMID:98332770; PMID:9665876
A;Accession: B71274
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-179 <COL>
A;Cross-references: UNIPROT:Q83822; GB:AE001255; GB:AE000520; NID:g3323156; PIDN:AAC6581
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0850
C;Superfamily: translation initiation factor IF-3

Query Match          91.7%; Score 22; DB 2; Length 179;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
    ||||
Db 155 GYSVE 159

```

## RESULT 36

D72459

hypothetical protein APE2320 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C;Accession: D72459

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A;Reference number: A72450; PMID:99310339; PMID:10382966

A;Accession: D72459

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-179 &lt;KAW&gt;

A;Cross-references: UNIPROT:Q9Y9G8; DBJ:AP000064; NID:g5105945; PIDN:BAA81332.1; PID:d1

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE2320

C;Superfamily: Aeropyrum pernix hypothetical protein APE2320

Query Match 91.7%; Score 22; DB 2; Length 179;

Best Local Similarity 80.0%; Pred. No. 2.9e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

||||

Db 145 GYAVE 149

## RESULT 37

T34851

probable secreted protein - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

C;Accession: T34851

R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, February 1999

A;Reference number: Z21559

A;Accession: T34851

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-180 &lt;OLI&gt;

A;Cross-references: UNIPROT:Q9Z5A5; EMBL:AL035478; PIDN:CAB36605.1; GSPDB:GN00070; SCOED

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SC0DB:SC2G5.18c

C;Superfamily: Streptomyces coelicolor probable secreted protein SC8A6.12c

Query Match 91.7%; Score 22; DB 2; Length 180;

Best Local Similarity 80.0%; Pred. No. 3e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

||||

Db 57 GYAVE 61

## RESULT 38

T06718

hypothetical protein T29H11.240 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T06718

R;Quetier, F.; Choisme, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigue

submitted to the Protein Sequence Database, April 1999

A;Reference number: Z15793

A;Accession: T06718

A;Molecule type: DNA

A;Residues: 1-180 &lt;QUE&gt;

A;Cross-references: UNIPROT:Q9STK4; EMBL:AL049659; GSPDB:GN00061; ATSP:T29H11.240

A;Experimental source: cultivar Columbia; BAC clone T29H11  
C;Genetics:  
A;Gene: ATGP:T29H11.240  
A;Map position: 3  
A;Introns: 53/1

Query Match 91.7%; Score 22; DB 2; Length 180;  
Best Local Similarity 80.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVE 5  
|||  
Db 63 GYSVE 67

RESULT 39  
E87289  
conserved hypothetical protein CC0326 [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: E87289  
R;Nierman, M.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: AB7249; MUID:21173698; PMID:11259647  
A;Accession: E87289  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-182 <STO>  
A;Cross-references: UNIPROT:Q9ABA5; GB:AE005673; NID:gl3421473; PIDN:AAK22313.1; GSPDB:G  
C;Genetics:  
A;Gene: CC0326

Query Match 91.7%; Score 22; DB 2; Length 182;  
Best Local Similarity 80.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVE 5  
|||  
Db 107 GYAVE 111

RESULT 40  
AG2528  
hypothetical protein alr7514 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120be  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AG2528  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takarawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AG2528  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-187 <KUR>  
A;Cross-references: UNIPROT:Q82SJ6; GB:AP003602; PIDN:BAE771157.1; PID:gl7134599; GSPDB:G  
C;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: alr7514  
A;Genome: plasmid

Query Match 91.7%; Score 22; DB 2; Length 187;  
Best Local Similarity 80.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 66 GYAVE 70

RESULT 41  
D87705  
intracellular septation protein A, probable [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: D87705  
R;Nierman, M.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: AB7249; MUID:21173698; PMID:11259647  
A;Accession: D87705  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-188 <STO>  
A;Cross-references: UNIPROT:Q9A288; GB:AE005673; NID:gl3425438; PIDN:AAK25640.1; GSPDB:G  
C;Genetics:  
A;Gene: CC3678

Query Match 91.7%; Score 22; DB 2; Length 188;  
Best Local Similarity 80.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVE 5  
|||  
Db 41 GYAVE 45

RESULT 42  
AF1504  
amidotransferases homolog hisH [imported] - Listeria innocua (strain Clip11262)  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AF1504  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AF1504  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-208 <GLA>  
A;Cross-references: UNIPROT:Q92E86; GB:AL592022; PIDN:CAC95806.1; PID:gl6413014; GSPDB:G  
A;Experimental source: strain Clip11262  
C;Genetics:  
A;Gene: hisH  
C;Superfamily: amidotransferase hisH; trpG homology

Query Match 91.7%; Score 22; DB 2; Length 208;  
Best Local Similarity 80.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVE 5  
|||  
Db 161 GYSVE 165

RESULT 43  
A41030  
troponin I, cardiac muscle - quail  
C;Species: Coturnix coturnix (quail)  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 13-Aug-1999  
C;Accession: A41030  
R;Hastings, K.E.M.; Koppe, R.I.; Marmor, B.; Bader, D.; Shimada, Y.; Toyota, N.  
J. Biol. Chem. 266, 19659-19665, 1991  
A;Title: Structure and developmental expression of troponin I isoforms. cDNA clone analysis



A:Reference number: A41030; MUID:920111768; PMID:1918073

A:Accession: A41030

A:Molecule type: mRNA

A:Residues: 1-208 <HAS>

A:Cross-references: GB:M73702; NID:g213647; PIDN:AAA49513.1; PID:g213648

C:Superfamily: troponin I

C:Keywords: actin binding; cardiac muscle; heart

Query Match 91.7%; Score 22; DB 2; Length 208;  
Best Local Similarity 80.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

|||

Db 22 GYAVE 26

#### RESULT 44

AH2716 NTP pyrophosphohydrolase, MutT family [imported] - Agrobacterium tumefaciens (strain C58)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004

C:Accession: AH2716

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AH2716

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-209 <KUR>

A:Cross-references: UNIPROT:Q8UGA2; GB:AE008688; PIDN:AAL42150.1; PID:g17739536; GSPDB:C

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atul137

A:Map position: circular chromosome

C:Superfamily: yffh protein; mutT domain homology

Query Match 91.7%; Score 22; DB 2; Length 209;  
Best Local Similarity 80.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

|||

Db 118 GYAVE 122

#### RESULT 45

A54984 EUF-1 protein precursor - mouse

N:Alternate names: Cek7 ligand

C:Species: Mus musculus (house mouse)

C:Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A54984; A55873

R:Cheng, H.J.; Flanagan, J.G.

Cell 79, 157-168, 1994

A:Title: Identification and cloning of ELF-1, a developmentally expressed ligand for the

A:Reference number: A54984; MUID:95007776; PMID:7522971

A:Accession: A54984

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-209 <CHE>

A:Cross-references: UNIPROT:P52801; GB:U14941; NID:9558836; PIDN:AAAS3636.1; PID:9558837

R:Shao, H.; Lou, L.; Pandey, A.; Verderame, M.F.; Siever, D.A.; Dixit, V.M.

J. Biol. Chem. 270, 3467-3470, 1995

A:Title: cDNA cloning and characterization of a Cek7 receptor protein-tyrosine kinase li

A:Reference number: A55873; MUID:95181289; PMID:7876076

A:Accession: A55873

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-209 <SHA>

A:Cross-references: GB:U14752; NID:g681886; PIDN:AAA68520.1; PID:g681887

C:Superfamily: axon guidance signal protein

C:Keywords: lipoprotein; membrane protein

Query Match 91.7%; Score 22; DB 2; Length 209;  
Best Local Similarity 80.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

|||

Db 54 GYTVE 58

#### RESULT 46

T32354

hypothetical protein C08E3.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T32354

R:Miller, N.; Kramer, J.; Keppler, D.

submitted to the EMBL Data Library, September 1997

A:Description: The sequence of C. elegans cosmid C08E3.

A:Reference number: Z21155

A:Accession: T32354

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-211 <MII>

A:Cross-references: UNIPROT:O17194; EMBL:AF025457; PIDN:AAB70965.1; GSPDB:GN000020; CESP:

A:Experimental source: strain Bristol N2; clone C08E3

C:Genetics:

A:Gene: CESP:C08E3.4

A:Map position: 2

A:Introns: 36/3; 149/3

Query Match 91.7%; Score 22; DB 2; Length 211;

Best Local Similarity 80.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

|||

Db 207 GYTVE 211

#### RESULT 47

JE0322

ephraim-A2 - human

C:Species: Homo sapiens (man)

C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004

C:Accession: JE0322

R:Aasheim, H.; Pedoutour, F.; Grosgeorge, J.; Loqtenberg, T.

Biochem. Biophys. Res. Commun. 252, 378-382, 1998

A:Title: Cloning, chromosomal mapping, and tissue expression of the gene encoding the humar

A:Reference number: JE0322; MUID:99045414; PMID:9828538

A:Accession: JE0322

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-213 <AAS>

A:Cross-references: UNIPROT:O43921; GB:AJ007292; NID:g3688367; PIDN:CAA07435.1; PID:g368

C:Superfamily: axon guidance signal protein

Query Match 91.7%; Score 22; DB 2; Length 213;

Best Local Similarity 80.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

|||

Db 58 GYTVE 62

#### RESULT 48

JQ2001

hypothetical 24K protein - lactic dehydrogenase virus  
 N:Alternate names: hypothetical 22.4K protein; ORF 5 protein  
 N:Contains: ORF 5 protein  
 C:Species: lactic dehydrogenase virus  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: J02001; J01994  
 R:Chen, Z.; Kuo, L.; Rowland, R.R.R.; Even, C.; Faaborg, K.S.; Plagemann, P.G.W.  
 J. Gen. Virol. 74, 643-650, 1993  
 A:Title: Sequences of 3' end of genome and of 5' end of open reading frame 1a of lactate  
 s.  
 A:Reference number: JQ1990; MUID:93224885; PMID:8385693  
 A:Accession: J02001  
 A:Molecule type: mRNA  
 A:Residues: 1-214 <CH>  
 A:Cross-references: UNIPROT:Q83022; GB:I06811  
 A:Experimental source: isolate P  
 A:Accession: JQ1994  
 A:Molecule type: mRNA  
 A:Residues: 16-214 <CH2>  
 A:Cross-references: GB:I06811  
 A:Experimental source: isolate P  
 C:Superfamily: lactic dehydrogenase virus hypothetical 24K protein  
 C:Keywords: envelope protein; glycoprotein  
 F:51,60,67/Binding site: carbohydrate (Asn) (covalent) #status Predicted

Query Match 91.7%; Score 22; DB 2; Length 214;  
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 76 GYAVE 80

RESULT 49  
 A72153  
 O2L protein - variola minor virus (strain Garcia-1966)  
 C:Species: variola minor virus  
 C:Date: 24-Nov-1999 #sequence\_revision 24-Nov-1999 #text\_change 15-Sep-2003  
 C:Accession: A72153  
 R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; Lopar  
 submitted to GenBank, March 1998  
 A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor  
 A:Reference number: A72150  
 A:Accession: A72153  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-220 <SHC>  
 A:Cross-references: GB:Y16780; NID:G5830555; PIDN:CAB54619.1; PID:G5830580  
 A:Experimental source: strain Garcia-1966  
 C:Genetics:  
 A:Gene: O2L  
 C:Superfamily: uncharacterized conserved protein

Query Match 91.7%; Score 22; DB 2; Length 220;  
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 15 GYSVE 19

RESULT 50  
 H42504  
 M2L protein - vaccinia virus (strain Copenhagen)  
 C:Species: vaccinia virus  
 A:Note: host Homo sapiens [man]  
 C:Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 09-Jul-2004  
 C:Accession: H42504  
 R:Johnson, G.P.  
 submitted to GenBank, June 1990  
 A:Reference number: A33172

A:Accession: H42504  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-220 <JOH>  
 A:Cross-references: UNIPROT:P21092  
 C:Superfamily: uncharacterized conserved protein

Query Match 91.7%; Score 22; DB 2; Length 220;  
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 15 GYSVE 19

RESULT 51  
 T28457  
 hypothetical protein O2L - variola major virus  
 C:Species: variola major virus  
 C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T28457  
 R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin,  
 Nature 366, 748-751, 1993  
 A:Title: Potential virulence determinants in terminal regions of variola smallpox virus  
 A:Reference number: Z20488; MUID:94088747; PMID:8264798  
 A:Accession: T28457  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-220 <MAS>  
 A:Cross-references: UNIPROT:P34017; EMBL:L22579; NID:G623595; PIDN:AAA60767.1; PID:G4389;  
 A:Experimental source: strain "Bangladesh-1975"  
 C:Superfamily: uncharacterized conserved protein

Query Match 91.7%; Score 22; DB 2; Length 220;  
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 15 GYSVE 19

RESULT 52  
 I36838  
 O2L protein - variola virus (strain India-1967)  
 C:Species: variola virus  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: I36838  
 R:Blinov, V.M.  
 submitted to GenBank, November 1992  
 A:Reference number: A36859  
 A:Accession: I36838  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-220 <BLI>  
 A:Cross-references: UNIPROT:P34017; GB:X69198; NID:G456758; PIDN:CAA48960.1; PID:G297201  
 C:Superfamily: uncharacterized conserved protein

Query Match 91.7%; Score 22; DB 2; Length 220;  
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 15 GYSVE 19

RESULT 53  
 C84383  
 hypothetical protein Vng2323h [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C;Accession: C94383  
 R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A.; Title: Genome sequence of Halobacterium species NRC-1.  
 A;Reference number: A84160; MUID:20504483; PMID:11016950  
 A;Accession: C94383  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-220 <STO>  
 A;Cross-references: UNIPROT:Q9HMZ0; GB:AE004437; NID:g10581735; PIDN:AAG20431.1; GSPDB:C  
 C;Genetics:  
 A;Gene: VNG2323H

Query Match 91.7%; Score 22; DB 2; Length 220;  
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 135 GYAVE 139

RESULT 54  
 F86930  
 probable two-component response regulator ML0174 [imported] - Mycobacterium leprae  
 C;Species: Mycobacterium leprae  
 C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C;Accession: F86930  
 R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holt, R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Squares, S.; Title: Massive gene decay in the leprosy bacillus.  
 A;Reference number: A86909; MUID:21128732; PMID:11234002  
 A;Accession: F86930  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-228 <STO>  
 A;Cross-references: UNIPROT:Q9CD68; GB:AL450380; NID:g13092534; PIDN:CAC29682.1; GSPDB:C  
 C;Genetics:  
 A;Gene: ML0174  
 C;Superfamily: ompR protein; response regulator homology

Query Match 91.7%; Score 22; DB 2; Length 228;  
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 24 GYSVE 28

RESULT 55  
 H64367  
 dolichyl-phosphate mannose synthase - Methanococcus jannaschii  
 C;Species: Methanococcus jannaschii  
 C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
 C;Accession: H64367  
 R;Bult, C.J.; Overbeek, R.; Kirkness, E.F.; Weinstein, K.G.; Merrick, J.M.; Glodek, A.; Raun, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.  
 A;Reference number: A64300; MUID:96337999; PMID:8688087  
 A;Accession: H64367  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-229 <BUL>

A;Cross-references: UNIPROT:Q57964; GB:U67504; NID:g1591248; PIDN:AAB98538.1;  
 C;Genetics:  
 A;Map position: RV482606-481917  
 C;Superfamily: Methanobacterium thermoautotrophicum dolichyl-phosphate mannose synthase 1

Query Match 91.7%; Score 22; DB 2; Length 229;  
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 175 GYAVE 179

RESULT 56  
 A70821  
 hypothetical protein Rv0981 - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C;Accession: A70821  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: A70821  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-230 <COL>  
 A;Cross-references: UNIPROT:O53894; GB:AL021999; GB:AL123456; NID:g3261538; PIDN:CAA17581  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 A;Gene: Rv0981  
 C;Superfamily: ompR protein; response regulator homology  
 C;Keywords: phosphoprotein  
 F:53/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 91.7%; Score 22; DB 2; Length 230;  
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 26 GYSVE 30

RESULT 57  
 S45089  
 hypothetical protein eta - Streptococcus pyogenes plasmid pDB101  
 C;Species: Streptococcus pyogenes  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
 C;Accession: S68123; S45089  
 R;Ceglowski, P.; Alonso, J.C.  
 Gene 145, 33-39, 1994  
 A;Title: Gene organization of the Streptococcus pyogenes plasmid pDB101.  
 A;Reference number: S68123; MUID:94320784; PMID:8093174  
 A;Accession: S68123  
 A;Molecule type: DNA  
 A;Residues: 1-231 <CE2>  
 A;Cross-references: UNIPROT:Q54941; EMBL:X66468; NID:g496500; PIDN:CAA47095.1; PID:G55991  
 A;Experimental source: plasmid pDB101  
 C;Genetics:  
 A;Genome: plasmid pDB101  
 A;Start codon: TTG

Query Match 91.7%; Score 22; DB 2; Length 231;  
 Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

```

Db          || ||
186 GYSVE 190

RESULT 58
G75075
hypothetical protein PAB1706 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: G75075
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: G75075
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-233 <KAW>
A;Cross-references: UNIPROT:Q9UZZ8; GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB4990
A;Experimental source: strain Orsay
C;Genetics:
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1634

Query Match          91.7%; Score 22; DB 2; Length 233;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
|| ||
Db 18 GYSVE 22

RESULT 59
AF2051
two-component system response regulator all1964 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Notes: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
R;Kanakako, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF2051
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-236 <KUR>
A;Cross-references: UNIPROT:Q8YVL1; GB:BA000019; PIDN:BA073663.1; PID:g17131054; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1964
C;Superfamily: ompR protein; response regulator homology

Query Match          91.7%; Score 22; DB 2; Length 236;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
|| ||
Db 30 GYTVE 34

RESULT 60
B84167
glucose-1-phosphate thymidyltransferase [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: B84167
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Lie
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: B84167
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-238 <STO>
A;Cross-references: UNIPROT:Q9HSV0; GB:AE004437; NID:g10579718; PIDN:AAG18702.1; GSPDB:GN
C;Genetics:
A;Gene: grad23

Query Match          91.7%; Score 22; DB 2; Length 238;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
|| ||
Db 206 GYSVE 210

RESULT 61
C82393
probable amino acid ABC transporter, periplasmic amino acid-binding protein VCA0978 [imp
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: C82393
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: C82393
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-243 <HBI>
A;Cross-references: UNIPROT:Q9KKX2; GB:AE004424; GB:AE003853; NID:g9658414; PIDN:AAF96874
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0978
A;Map position: 2

Query Match          91.7%; Score 22; DB 2; Length 243;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
|| ||
Db 56 GYTVE 60

RESULT 62
G75259
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: G75259
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567466
A;Accession: G75259
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-244 <WHI>
A;Cross-references: UNIPROT:Q9RRE4; GB:AE002084; GB:AE000513; NID:g6460372; PIDN:AAF1208;
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2548
A;Map position: 1

```

C:Superfamily: hypothetical protein MG332

Query Match 91.7%; Score 22; DB 2; Length 244;  
Best Local Similarity 80.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 190 GYAVE 194

RESULT 63

T45446 probable two-component response regulator [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T45446

R:James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, February 1998

A:Reference number: Z22967

A:Status: preliminary

A:Note: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-253 <JAM>

A:Cross-references: UNIPROT:Q9Z5G8; EMBL:AL035500; PIDN:CAB36688.1

A:Experimental source: cosmid L373

C:Genetics:

A:Note: MLCB373.26

C:Superfamily: ompR protein; response regulator homology

Query Match 91.7%; Score 22; DB 2; Length 253;  
Best Local Similarity 80.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 49 GYSVE 53

RESULT 64

C97498 hypothetical 21.7K protein in tktB-narQ intergenic region [imported] - Agrobacterium tum

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004

C:Accession: C97498

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-254 <KUR>

A:Cross-references: UNIPROT:Q8UGA2; GB:AE007869; PIDN:AAK86940.1; PID:g15156172; GSPDB:G

C:Genetics:

A:Gene: AGR\_C 2106

A:Map position: circular chromosome

Query Match 91.7%; Score 22; DB 2; Length 254;  
Best Local Similarity 80.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 163 GYAVE 167

RESULT 65

GB4279

stationary phase survival protein [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: G84279

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;

J.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Lia

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: G84279

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-258 <STO>

A:Cross-references: UNIPROT:Q9HQB2; GB:AE004437; NID:g10580771; PIDN:AAG19603.1; GSPDB:G

C:Genetics:

A:Gene: surE

C:Superfamily: stationary-phase survival protein SurE

Query Match 91.7%; Score 22; DB 2; Length 258;  
Best Local Similarity 80.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 62 GYAVE 66

RESULT 66

F70432

conserved hypothetical protein aq\_1528 - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004

C:Accession: F70432

R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove

V. Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: F70432

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-270 <AOF>

A:Cross-references: UNIPROT:O67488; GB:AE000744; NID:g2983891; PIDN:AAC07456.1; PID:g2983

A:Experimental source: strain VF5

C:Genetics:

A:Gene: aq\_1528

Query Match 91.7%; Score 22; DB 2; Length 270;  
Best Local Similarity 80.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 22 GYTVS 26

RESULT 67

B95852

conserved hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplas

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004

C:Accession: B95852

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmaster, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endos

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: B95852

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-272 <KUR>

A:Cross-references: UNIPROT:Q92X80; GB:AL591985; PIDN:CAC48482.1; PID:g15139954; GSPDB:G

A:Experimental source: strain 1021, megaplasmid pSymb

R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:References: number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: SMB20082  
A:Genome: plasmid

Query Match 91.7%; Score 22; DB 2; Length 272;  
Best Local Similarity 80.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 86 GYAVE 90

RESULT 68  
WMBP8H  
gene 8.5 protein - phage PZA  
C:Species: phage PZA  
A:Note: host Bacillus subtilis  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 09-Jul-2004  
C:Accession: C24831  
R:Paces, V.; Vlcek, C.; Urbanek, P.  
Gene 44, 107-114, 1986  
A:Title: Nucleotide sequence of the late region of Bacillus subtilis phage PZA, a close  
A:Reference number: A91550; MUID:87031573; PMID:3095188  
A:Accession: C24831  
A:Molecule type: DNA  
A:Residues: 1-280 <PAC>  
A:Cross-references: UNIPROT:P07532; GB:M11813; GB:M13904; GB:M13905; NID:G216046; PIDN:A  
A:Note: the authors translated the codon CTT for residue 207 as Pro, GGG for residue 208  
2 as Val, and GCG for residue 213 as Ser  
C:Genetics:  
A:Gene: 8.5  
C:Superfamily: phage Nf assembly protein  
C:Keywords: capsid assembly; late protein

Query Match 91.7%; Score 22; DB 1; Length 280;  
Best Local Similarity 80.0%; Pred. No. 4.7e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 96 GYAVE 100

RESULT 69  
WMBP8H  
Gene 8.5 protein - phage phi-29  
N:Alternate names: assembly protein; head fiber protein  
C:Species: phage phi-29  
A:Note: host Bacillus subtilis  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 09-Jul-2004  
C:Accession: C25816  
R:Vlcek, C.; Paces, V.  
Gene 46, 215-225, 1986  
A:Title: Nucleotide sequence of the late region of Bacillus phage phi-29 completes the  
A:Reference number: A25816; MUID:87106857; PMID:3803926  
A:Accession: C25816  
A:Molecule type: DNA  
A:Residues: 1-280 <VLC>  
A:Cross-references: UNIPROT:P20344; GB:M14782; NID:G215323; PIDN:AAA32281.1; PID:G215326  
C:Genetics:  
A:Gene: 8.5  
C:Superfamily: phage Nf assembly protein  
C:Keywords: capsid assembly; head protein; late protein

Query Match 91.7%; Score 22; DB 1; Length 280;

Best Local Similarity 80.0%; Pred. No. 4.7e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 96 GYAVE 100

RESULT 70  
T41846  
HE65 orf105 - Bombyx mori nuclear polyhedrosis virus (isolate T3)  
C:Species: Bombyx mori nuclear polyhedrosis virus, BMSNPV  
A:Variety: isolate T3  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T41846  
R:Gomi, S.; Majima, K.; Maeda, S.  
J. Gen. Virol. 80, 1323-1337, 1999  
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.  
A:Reference number: Z22020; MUID:99281911; PMID:10355780  
A:Accession: T41846  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-289 <KAM>  
A:Cross-references: UNIPROT:O92465; EMBL:L33180; PIDN:AA63775.1  
A:Experimental source: isolate T3  
C:Genetics:  
A:Note: he65

Query Match 91.7%; Score 22; DB 2; Length 289;  
Best Local Similarity 80.0%; Pred. No. 4.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 14 GYAVE 18

RESULT 71  
T27430  
hypothetical protein Y79H2A.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T27430  
R:Matthews, L.  
submitted to the EMBL Data Library, September 1999  
A:Reference number: Z20365  
A:Accession: T27430  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-293 <WIL>  
A:Cross-references: UNIPROT:Q9UIR9; EMBL:AL110501; NID:e1542357; PIDN:CAB54508.1; CESP:Y:  
A:Experimental source: clone Y79H2A  
C:Genetics:  
A:Gene: CESP:Y79H2A.2  
A:Introns: 30/3; 50/1; 85/1; 112/3; 180/3; 223/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein Y79H2A.2

Query Match 91.7%; Score 22; DB 2; Length 293;  
Best Local Similarity 80.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 276 GYSVE 280

RESULT 72  
S55795  
arginase (EC 3.5.3.1) rocF - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S55795; B69694; S49269  
R:Gardan, R.; Rapoport, G.; Debarbouille, M.

```

Qy      1 GYXVE 5
      ||||
Db      61 GYSVE 65

RESULT 74
A70441
glycerol-3-phosphate dehydrogenase (NAD) (EC 1.1.1.8) - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 16-Aug-2004
C;Accession: A70441
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Overman, J.A.
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: A70441
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-313 <AQF>
A;Cross-references: GB:AE000748; NID:g2983960; PIDN:AAC07511.1; PID:g2983964; GB:AE00065;
A;Experimental source: strain VF5
C;Genetics:
A;Gene: gspA
C;Superfamily: Glycerol-3-phosphate dehydrogenase (NAD)
C;Keywords: oxidoreductase

Query Match      91.7%; Score 22; DB 2; Length 313;
Best Local Similarity 80.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      ||||
Db      246 GYSVE 250

RESULT 75
D72484
probable ATP-binding protein APE2516 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 16-Aug-2004
C;Accession: D72484
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kikuchi, A.
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix K1
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: D72484
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-323 <KAW>
A;Cross-references: UNIPROT:Q9Y8W8; DDBJ:AP000064; NID:g5105945; PIDN:BA081532.1; PID:di00000000
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2516
C;Superfamily: ATP-binding cassette homology
F;29-221/Domain: ATP-binding cassette homology <ABC>

Query Match      91.7%; Score 22; DB 2; Length 323;
Best Local Similarity 80.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      ||||
Db      2 GYAVE 6

Search completed: November 1, 2004, 21:35:11
Job time : 35 secs

```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:15:12 ; Search time 127.5 Seconds

(without alignment)  
27.076 Million cell updates/sec

Title: US-10-030-194A-6

Perfect score: 24

Sequence: 1 GYVEX 6

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	22	91.7	47	2	Q8W6Q0	bacterioph
2	22	91.7	51	2	Q8LLV4	Q8llv4 ipomoea bat
3	22	91.7	76	2	Q8U8S6	Q8u8s6 agrobacteri
4	22	91.7	76	2	Q8XTM4	Q8xtm4 ralstonia s
5	22	91.7	78	1	DY21_ECOLI	P00383 escherichia
6	22	91.7	78	1	DY22_ECOLI	P00384 escherichia
7	22	91.7	78	1	DY23_ECOLI	P05794 escherichia
8	22	91.7	78	2	Q71RY0	Q71ry0 aeromonas s
9	22	91.7	78	2	Q79NS3	Q79ns3 uncultured
10	22	91.7	78	2	Q8VNN1	Q8vnn1 escherichia
11	22	91.7	78	2	Q6D6P6	Q6d6p6 erwinia car
12	22	91.7	78	2	AAH89276	AAh89276 escherich
13	22	91.7	78	2	AAH82255	AAh82255 escherich
14	22	91.7	78	2	AAK53555	AAk53555 aeromonas
15	22	91.7	78	2	AAK95982	AAk95982 escherich
16	22	91.7	78	2	AAH41415	AAh41415 gamma-pro
17	22	91.7	81	2	Q89013	Q89013 proteobact
18	22	91.7	82	1	ATPH_CHLRE	Q37304 chlamydomon
19	22	91.7	82	1	ATPH_CHLVU	P56297 chlorella v
20	22	91.7	83	2	Q8PTB1	Q8ptb1 methanosarc
21	22	91.7	84	2	Q8Z2J8	Q8z2j8 pyrobaculum
22	22	91.7	86	1	YEB6_CLOB	Q97j00 clostridium
23	22	91.7	86	2	Q8KK08	Q8kk08 proteus vul
24	22	91.7	87	2	Q6D329	Q6d329 erwinia car
25	22	91.7	88	2	Q8TPE9	Q8tpe9 methanosarc
26	22	91.7	91	2	Q48265	Q48265 haemophilus
27	22	91.7	94	2	Q8Z2J2	Q8z2j2 rhizobium i
28	22	91.7	97	1	YXVB_CALSA	P23554 caldocellum
29	22	91.7	97	2	Q8GL63	Q8gl63 uncultured
30	22	91.7	99	2	Q8IKH6	Q8ikh6 plasmodium
31	22	91.7	101	2	Q7MQT5	Q7mqt5 wolinnella s

32	22	91.7	108	1	YGSN_YEAST	P40326 saccharomyc
33	22	91.7	108	2	Q7Y3H2	Q7y3h2 enterobacte
34	22	91.7	108	2	Q6FPA4	Q6ffa4 acinetobact
35	22	91.7	108	2	AAS56709	Aas56709 saccharom
36	22	91.7	109	2	O58646	O58646 pyrococcus
37	22	91.7	112	2	Q7PFW9	Q7pfv9 anopheles g
38	22	91.7	114	2	O6N3A0	O6n3a0 rhodopsu
39	22	91.7	114	2	CAE29235	CAe29235 rhodopseu
40	22	91.7	115	2	Q979N5	Q979n5 thermoplas
41	22	91.7	115	2	Q84XP9	Q84xp9 brassica ra
42	22	91.7	119	2	Q7CTV2	Q7ctv2 agrobacteri
43	22	91.7	120	2	O28647	O28647 archaeglob
44	22	91.7	122	2	Q9AAK1	Q9aak1 caulobacter
45	22	91.7	123	2	Q9C7K1	Q9c7k1 arabisopsis
46	22	91.7	127	2	Q6NME7	Q6nme7 arabidopsi
47	22	91.7	127	2	AAS49076	Aas49076 arabidops
48	22	91.7	129	2	Q14556	Q14556 homo sapien
49	22	91.7	131	1	RS6E_HALN1	Rshmj5 halobacteri
50	22	91.7	133	2	Q8C9X7	Q8c9x7 mus musculu
51	22	91.7	134	2	Q9LAV0	Q9lav0 caulobacter
52	22	91.7	134	2	Q7V168	Q7v168 prochloroco
53	22	91.7	134	2	Q9CYV3	Q9cyv3 m mus muscu
54	22	91.7	135	2	Q853T4	Q853t4 mycobacteri
55	22	91.7	136	1	RUVX_DEIRA	Q9tri2 deinococcus
56	22	91.7	136	2	Q9HYI5	Q9hyi5 pseudomonas
57	22	91.7	137	1	YH92_ARCFU	YH9228 archaeglob
58	22	91.7	139	2	Q9KHT5	Q9kht5 pseudomonas
59	22	91.7	139	2	Q88ME0	Q88me0 pseudomonas
60	22	91.7	142	2	Q9Y9Z8	Q9y9z8 aeropyrum p
61	22	91.7	144	2	Q7S0W9	Q7s0w9 neurospora
62	22	91.7	145	2	Q88E22	Q88e22 pseudomonas
63	22	91.7	146	1	PFDA_PYRFU	PFda pyrococcus
64	22	91.7	148	1	PFDA_PYRAB	PFda pyrococcus
65	22	91.7	148	1	PFDA_PYRHO	O58263 pyrococcus
66	22	91.7	148	1	RL9_LISIN	Q92fq7 listeria in
67	22	91.7	148	1	RL9_LISMO	Q8var2 listeria mo
68	22	91.7	148	2	Q725B1	Q725b1 listeria mo
69	22	91.7	148	2	AAO2850	AAo2850 listeria
70	22	91.7	150	2	Q8GPS7	Q8gps7 pseudomonas
71	22	91.7	151	2	Q6V4S8	Q6v4s8 enterococc
72	22	91.7	151	2	AAH10433	AAh10433 enterococ
73	22	91.7	152	1	YB83_METJA	Yb83 methanococc
74	22	91.7	153	2	Q7W0H6	Q7w0h6 bordetella
75	22	91.7	153	2	Q7W3N1	Q7w3n1 bordetella
76	22	91.7	153	2	Q7WF07	Q7wf07 bordetella
77	22	91.7	153	2	Q9A023	Q9a023 streptococc
78	22	91.7	154	2	Q59820	Q59820 staphylococ
79	22	91.7	154	2	Q8DS13	Q8ds13 streptococc
80	22	91.7	155	2	Q82YZ6	Q82yz6 enterococcu
81	22	91.7	157	2	Q52245	Q52245 enterococcu
82	22	91.7	158	2	Q91233	Q91233 pseudomonas
83	22	91.7	162	1	YK41_ARCFU	Yk41 archaeglob
84	22	91.7	165	2	Q6MPK7	Q6mpk7 bdellovibri
85	22	91.7	165	2	CAE78790	CAe78790 bdellovib
86	22	91.7	167	2	Q6E432	Q6e432 nootka lupi
87	22	91.7	168	2	O87846	O87846 streptomyce
88	22	91.7	168	2	Q897V6	Q897v6 clostridium
89	22	91.7	170	2	Q8ZYH3	Q8zyh3 pyrobaculum
90	22	91.7	171	2	Q7W2M0	Q7w2m0 bordetella
91	22	91.7	171	2	Q7WDL6	Q7wdl6 bordetella
92	22	91.7	171	2	Q815R0	Q815r0 bacillus ce
93	22	91.7	174	1	YN11_MYCTU	YN11 mycobacteri
94	22	91.7	174	2	O67188	O67188 aquifex aeo
95	22	91.7	176	2	Q74G14	Q74g14 geobacter s
96	22	91.7	176	2	AAH33770	AAh33770 geobacter
97	22	91.7	178	2	O33538	O33538 rhodococcus
98	22	91.7	179	1	IF3_TREPA	IF33822 treponema p
99	22	91.7	179	2	Q9Y9G8	Q9y9g8 aeropyrum p
100	22	91.7	180	2	Q9STK4	Q9stk4 arabidopsi
101	22	91.7	180	2	Q9Z5A5	Q9z5a5 streptomyce
102	22	91.7	180	2	Q98AL6	Q98al6 rhizobium l
103	22	91.7	180	2	AAH41790	AAh41790 arabidops
104	22	91.7	181	2	Q9AYX9	Q9ayx9 lactococcus

105	22	91.7	182	2	Q9AB5	Q9ab5 caulobacter	178	22	91.7	232	2	Q6L2N4	Q6lzn4 methanococ
106	22	91.7	187	2	Q8ZS06	Q8zsj6 anabaena sp	179	22	91.7	232	2	Q7XYK9	Q7xyk9 chlorarachn
107	22	91.7	188	1	ISPZ_CAUCR	Q9a288 caulobacter	180	22	91.7	232	2	CAF30146	CAF30146 methanoco
108	22	91.7	191	2	OS3055	Q53055 lactococcus	181	22	91.7	233	2	Q9UZZ8	Q9uzz8 pyrococcus
109	22	91.7	191	2	Q7ZU20	Q7zu20 brachydanio	182	22	91.7	233	2	Q9KJX2	Q9kix2 pyrococcus
110	22	91.7	191	2	AAH71487	Aah71487 brachydan	183	22	91.7	233	2	Q8Q0W0	Q8qow0 porcine cir
111	22	91.7	195	2	O17194	O17194 caenorhabdi	184	22	91.7	233	2	Q8ZWB0	Q8zwb0 pyrobaculum
112	22	91.7	195	2	Q6XUN3	Q6xum3 pseudomonas	185	22	91.7	235	2	Q8NU43	Q8nu43 homo sapien
113	22	91.7	195	2	AA444262	Aap44262 pseudomon	186	22	91.7	236	2	Q8YVL1	Q8yvl1 anabaena sp
114	22	91.7	198	2	Q6JCV9	Q6jcv9 xanthomonas	187	22	91.7	237	2	Q9MC25	Q9mc25 lactococcus
115	22	91.7	198	2	Q8PDX9	Q8pdx9 xanthomonas	188	22	91.7	238	2	Q8HSV0	Q8hsv0 halobacteri
116	22	91.7	198	2	Q8PQU8	Q8pqu8 xanthomonas	189	22	91.7	238	2	Q8TYE7	Q8tye7 methanopyru
117	22	91.7	198	2	Q8EM62	Q8em62 oceanobacil	190	22	91.7	240	2	Q8DAC3	Q8dac3 mus musculu
118	22	91.7	199	2	Q83022	Q83022 lactate deh	191	22	91.7	242	2	Q81260	Q81260 tripsacum d
119	22	91.7	199	2	Q83032	Q83032 lactate deh	192	22	91.7	242	2	Q74FM9	Q74fm9 geobacter s
120	22	91.7	199	2	Q8YVS38	Q8yvs38 lactate deh	193	22	91.7	243	2	Q8KKX2	Q8kkx2 vibrio chol
121	22	91.7	199	2	Q8YVS38	Q8yvs38 lactate deh	194	22	91.7	243	2	Q9KX2	Q9kx2 pyrobaculum
122	22	91.7	199	2	Q8PXD0	Q8pxd0 lactate deh	195	22	91.7	246	2	Q879L2	Q8t9l2 drosophila
123	22	91.7	200	2	Q82048	Q82048 plasmid phv	196	22	91.7	246	2	Q87832	Q87832 streptomyc
124	22	91.7	200	2	Q83AC9	Q83ac9 coxiella bu	197	22	91.7	246	2	Q6FVY7	Q6fyv7 bartonella
125	22	91.7	201	2	Q83AC9	Q83ac9 corynebacte	198	22	91.7	246	2	Q8ZVL3	Q8zvl3 pyrobaculum
126	22	91.7	201	2	CAF19028	CAF19028 corynebac	199	22	91.7	250	2	Q7X674	Q7x674 oryza sativ
127	22	91.7	206	2	Q76778	Q76778 haemopis ma	200	22	91.7	251	2	CAD41850	Cad41850 oryza sat
128	22	91.7	207	1	TRIC_COTUA	P27672 coturnix co	201	22	91.7	251	2	CAE05963	CAe05963 oryza sat
129	22	91.7	208	1	HIS5_LISIN	Q92e86 listeria in	202	22	91.7	251	2	Q88964	Q88964 vaccinia vi
130	22	91.7	208	2	Q6S7R4	Q6s7r4 meleagris g	203	22	91.7	252	2	Q88964	Q88964 vaccinia vi
131	22	91.7	208	2	Q6S7R6	Q6s7r6 meleagris g	204	22	91.7	252	2	Q88964	Q88964 vaccinia vi
132	22	91.7	208	2	Q6S7R6	Q6s7r6 meleagris g	205	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
133	22	91.7	208	2	AA545403	AA545403 gallus ga	206	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
134	22	91.7	208	2	AA545404	AA545404 meleagris	207	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
135	22	91.7	208	2	AA545405	AA545405 meleagris	208	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
136	22	91.7	209	1	EPA2_MOUSE	P52801 mus musculu	209	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
137	22	91.7	209	2	Q8UGA2	Q8uga2 agrobacteri	210	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
138	22	91.7	210	2	Q8AZ77	Q8az77 lactococcus	211	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
139	22	91.7	211	2	Q8DWJ9	Q8dwj9 streptococc	212	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
140	22	91.7	213	1	EPA2_HUMAN	Q43921 homo sapien	213	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
141	22	91.7	213	1	Q8105	Q8105 bacterioph	214	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
142	22	91.7	214	2	Q8HG50	Q8hg50 verticilliu	215	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
143	22	91.7	214	2	Q8HG50	Q8hg50 verticilliu	216	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
144	22	91.7	215	2	Q8E501	Q8e501 lactate deh	217	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
145	22	91.7	215	2	Q8E501	Q8e501 lactate deh	218	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
146	22	91.7	219	2	Q8RGL8	Q8rgl8 fusbacteri	219	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
147	22	91.7	220	1	VM02_VACCC	P21092 vaccinia vi	220	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
148	22	91.7	220	1	VM02_VARV	P34017 variola vir	221	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
149	22	91.7	220	2	Q8E908	Q8e908 halobacteri	222	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
150	22	91.7	220	2	Q6RZR8	Q6rzzr shewanella	223	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
151	22	91.7	220	2	Q6RZR8	Q6rzzr rabbitpox v	224	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
152	22	91.7	220	2	Q76Q55	Q76q55 variola min	225	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
153	22	91.7	220	2	Q80HY2	Q80hy2 vaccinia vi	226	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
154	22	91.7	220	2	Q89082	Q89082 variola vir	227	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
155	22	91.7	220	2	Q8V554	Q8v554 variola vir	228	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
156	22	91.7	220	2	AA549736	AA549736 rabbitpox v	229	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
157	22	91.7	223	2	Q6NE98	Q6ne98 corynebacte	230	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
158	22	91.7	223	2	CAE50900	CAe50900 corynebac	231	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
159	22	91.7	224	2	Q6FZA8	Q6fza8 bartonella	232	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
160	22	91.7	224	2	Q6G2V0	Q6g2v0 bartonella	233	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
161	22	91.7	225	2	Q8QRU8	Q8gru8 chimpanzee	234	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
162	22	91.7	226	2	Q8ACD5	Q8acd5 vibrio salm	235	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
163	22	91.7	227	2	Q8PSW7	Q8psw7 xanthomonas	236	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
164	22	91.7	228	2	Q742C1	Q742c1 mycobacteri	237	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
165	22	91.7	228	2	Q9CD68	Q9cd68 mycobacteri	238	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
166	22	91.7	228	2	Q9QBW9	Q9qbw9 lily latent	239	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
167	22	91.7	228	2	AA503233	AA503233 mycobacte	240	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
168	22	91.7	229	1	Y544_METJA	Q57964 methanococ	241	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
169	22	91.7	229	2	Q6DS06	Q6d9u6 erwinia car	242	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
170	22	91.7	230	2	Q84BX0	Q84bx0 mycobacteri	243	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
171	22	91.7	230	2	Q53894	Q53894 mycobacteri	244	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
172	22	91.7	230	2	Q7UUX4	Q7uux4 mycobacteri	245	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
173	22	91.7	231	2	Q54941	Q54941 streptococc	246	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
174	22	91.7	231	2	Q9AL06	Q9al06 enterococcu	247	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
175	22	91.7	231	2	Q7NSF1	Q7nsf1 photorhabdu	248	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
176	22	91.7	231	2	AA272211	AA272211 streptoco	249	22	91.7	253	2	Q925G8	Q925g8 mycobacteri
177	22	91.7	231	2	AA272211	AA272211 streptoco	250	22	91.7	253	2	Q925G8	Q925g8 mycobacteri

251	22	91.7	281	2	AAT31421	Aat31421 bacillus	324	22	91.7	338	2	Q7QFT3	Q7qft3 anopheles g
252	22	91.7	282	1	DAAA_STAHA	P54694 staphylococ	325	22	91.7	338	2	Q85729	Q85729 streptomyce
253	22	91.7	283	1	EFTS_BIFLO	Q8g485 bifidobacte	326	22	91.7	338	2	P97061	P97061 streptomyce
254	22	91.7	284	2	Q21311	Q21311 tetrahymena	327	22	91.7	339	2	Q8TR96	Q8tr96 methanosarc
255	22	91.7	284	2	Q9T7M5	Q9t7m5 tetrahymena	328	22	91.7	340	1	CBID_PYRAE	Q8z2b0 pyrobaculum
256	22	91.7	284	2	Q950Y3	Q950y3 tetrahymena	329	22	91.7	341	2	Q735G4	Q735g4 bacillus ce
257	22	91.7	284	2	Q8FM88	Q8fm88 corynebacte	330	22	91.7	341	2	AA42098	AA42098 bacillus
258	22	91.7	287	2	Q9JPG1	Q9jpg1 neisseria m	331	22	91.7	342	2	Q8PJG0	Q8pjg0 xanthomonas
259	22	91.7	287	2	Q9JPI7	Q9jpi7 neisseria m	332	22	91.7	342	2	Q8XJF3	Q8xjf3 clostridium
260	22	91.7	289	2	Q8PX21	Q8px21 methanosarc	333	22	91.7	343	2	Q7NS68	Q7ns68 chromobacte
261	22	91.7	289	2	Q924G5	Q924g5 bombyx mori	334	22	91.7	343	2	Q98HF4	Q98hf4 rhizobium l
262	22	91.7	292	2	Q8SQS0	Q8sgs0 encephalito	335	22	91.7	344	1	SYFA_BACSU	P17921 bacillus su
263	22	91.7	292	2	Q7X347	Q7x347 uncultured	336	22	91.7	344	2	Q916H0	Q916h0 pseudomonas
264	22	91.7	293	2	Q9UIR9	Q9uir9 caenorhabdi	337	22	91.7	345	1	GBA4_DICDI	P34042 dictyosteli
265	22	91.7	293	2	Q6J5H7	Q6j5h7 haemophilus	338	22	91.7	345	2	Q286G5	Q286g5 archaeoglob
266	22	91.7	294	2	Q89ZRO	Q89zro bacteroides	339	22	91.7	345	2	Q6W295	Q6w295 rhizobium s
267	22	91.7	296	1	ARGI_BACSU	P39138 bacillus su	340	22	91.7	345	2	Q92NH1	Q92nh1 rhizobium m
268	22	91.7	296	2	Q8PKA9	Q8pkag xanthomonas	341	22	91.7	345	2	AAQ87122	AAq87122 rhizobium
269	22	91.7	297	2	Q9H8G7	Q9h8g7 heterodoxus	342	22	91.7	346	2	Q6JBI3	Q6jbi3 dictyocaulu
270	22	91.7	299	2	Q9HIS6	Q9hib6 thermoplasma	343	22	91.7	346	2	Q48229	Q48229 haemophilus
271	22	91.7	299	2	Q9HUA1	Q9hual pseudomonas	344	22	91.7	346	2	Q9I3T5	Q9i3t5 pseudomonas
272	22	91.7	300	2	Q72IK0	Q72ik0 thermus the	345	22	91.7	346	2	Q987Q0	Q987q0 rhizobium l
273	22	91.7	300	2	AA881474	AA881474 thermus t	346	22	91.7	346	2	AAT06310	Aat06310 dictyocau
274	22	91.7	301	2	Q75JN4	Q75jin4 dictyosteli	347	22	91.7	347	2	Q88C43	Q88c43 pseudomonas
275	22	91.7	301	2	AA538699	AA538699 dictyoste	348	22	91.7	348	2	Q8RX39	Q8rx39 arabidopsis
276	22	91.7	305	2	Q7UDB9	Q7ujb9 rhodopirell	349	22	91.7	349	1	GBT_XENLA	P38407 xenopus lae
277	22	91.7	305	2	Q98CV9	Q98cv9 rhizobium l	350	22	91.7	350	2	Q6MX97	Q6mx97 serratia ma
278	22	91.7	306	2	Q9L9Y3	Q9l9y3 lactococcus	351	22	91.7	350	2	CAE51827	CAe51827 serratia
279	22	91.7	308	2	Q7VYN9	Q7vyn9 bordetella	352	22	91.7	351	1	CBIH_METTH	O27454 methanobact
280	22	91.7	308	2	Q7W6H2	Q7w6h2 bordetella	353	22	91.7	352	2	Q8TWI9	Q8twi9 methanopyru
281	22	91.7	308	2	Q7WIE3	Q7wie3 bordetella	354	22	91.7	352	2	Q8VZJ1	Q8vzj1 arabidopsis
282	22	91.7	308	2	Q82ZN2	Q82zn2 enterococcu	355	22	91.7	352	2	Q8ZED6	Q8zed6 streptomyce
283	22	91.7	312	2	Q9L3H9	Q9l3h9 rhizobium l	356	22	91.7	354	2	Q6MPJ7	Q6mpj7 bdellovibri
284	22	91.7	314	2	Q8K2K0	Q8k2k0 mus musculu	357	22	91.7	354	2	Q83DQ3	Q83dq3 coxiella bu
285	22	91.7	315	2	Q8NTH1	Q8nth1 corynebacte	358	22	91.7	354	2	CAE78801	CAe78801 bdellovib
286	22	91.7	315	2	CAF19050	CAF19050 corynebac	359	22	91.7	356	1	CARA_THETN	O8rbk1 thermoeaer
287	22	91.7	319	2	Q75CN3	Q75cn3 ashbya goas	360	22	91.7	356	2	Q9UVX8	Q9uvx8 emericella
288	22	91.7	319	2	AA551114	AA551114 ashbya go	361	22	91.7	356	2	Q8TFX6	Q8tfx6 aspergillus
289	22	91.7	320	2	Q7QMA7	Q7qma7 anopheles g	362	22	91.7	356	2	Q918T3	Q918t3 human papil
290	22	91.7	321	2	Q8KNG0	Q8kng0 micromonosop	363	22	91.7	356	2	Q918T5	Q918t5 human papil
291	22	91.7	322	1	K6PE_STAEP	Q8cn68 staphylococ	364	22	91.7	356	2	CAF32008	CAF32008 aspergill
292	22	91.7	322	1	RIAB_TIPV	Q98vg9 f replicase	365	22	91.7	357	2	Q6Z2J7	Q6z2j7 oryza sativ
293	22	91.7	322	2	Q8LIC7	Q8lic7 enterococcu	366	22	91.7	357	2	BAD16143	Bad16143 oryza sat
294	22	91.7	323	2	Q9Y8W8	Q9y8w8 aeropyrum p	367	22	91.7	358	2	Q7Z9K5	Q7z9k5 penicillium
295	22	91.7	324	1	CATV_NPVCD	P41479 choristoneu	368	22	91.7	359	1	Y130_ARCFU	O30107 archaeoglob
296	22	91.7	324	1	CATV_NPVCF	P41715 choristoneu	369	22	91.7	359	2	Q918T0	Q918t0 human papil
297	22	91.7	324	1	CATV_NPVOP	O10364 oryza psau	370	22	91.7	360	1	CARA_STRP8	P58894 streptococc
298	22	91.7	324	1	GPDA_AQUAE	O67555 aquifex aeo	371	22	91.7	360	1	CARA_STRPY	Q940c7 streptococc
299	22	91.7	324	2	Q986J3	Q986j3 rhizobium l	372	22	91.7	361	2	Q840G9	Q840g9 actinobacil
300	22	91.7	324	2	Q6VTL7	Q6vtl7 choristoneu	373	22	91.7	361	2	Q9A2C3	Q9a2c3 caulobacter
301	22	91.7	324	2	AAQ91676	AAq91676 choristoneu	374	22	91.7	361	2	Q918T8	Q918t8 human papil
302	22	91.7	325	1	ODPB_STAAM	Q916h5 staphylococ	375	22	91.7	361	2	Q918U1	Q918u1 human papil
303	22	91.7	325	1	ODPB_STAEP	Q8cpn2 staphylococ	376	22	91.7	361	2	Q918U3	Q918u3 human papil
304	22	91.7	325	2	Q6GAC0	Q6gac0 staphylococ	377	22	91.7	361	2	Q918U5	Q918u5 human papil
305	22	91.7	325	2	Q6GHZ1	Q6ghz1 staphylococ	378	22	91.7	361	2	Q918U7	Q918u7 human papil
306	22	91.7	326	2	Q84HJ9	Q84hj9 streptomyce	379	22	91.7	362	2	Q99LL4	Q99ll4 mus musculu
307	22	91.7	326	2	Q6FAC8	Q6fac8 acinetobact	380	22	91.7	363	1	CARA_LISIN	Q92ah2 listeria in
308	22	91.7	327	2	Q8PSH1	Q8psh1 methanosarc	381	22	91.7	363	1	CARA_LISMO	Q8y664 listeria mo
309	22	91.7	329	2	Q7NN85	Q7nn85 gloebacter	382	22	91.7	363	2	Q9YD80	Q9y80 aeropyrum p
310	22	91.7	330	2	Q8TMU0	Q8tmu0 methanosarc	383	22	91.7	363	2	Q8L857	Q8l857 arabidopsis
311	22	91.7	331	2	Q948Z3	Q948z3 solanum tub	384	22	91.7	363	2	Q7IY10	Q7lyi0 listeria mo
312	22	91.7	332	2	Q91ER7	Q91er7 cydia pomon	385	22	91.7	363	2	AAT04634	Aat04634 listeria
313	22	91.7	333	2	Q30357	Q30357 neisseria g	386	22	91.7	365	1	VE2_HPV16	P03120 human papil
314	22	91.7	333	2	Q6MF42	Q6mf42 parachlamyd	387	22	91.7	365	2	Q9N061	Q9n061 macaca faec
315	22	91.7	333	2	CAF22807	CAF22807 parachlam	388	22	91.7	365	2	Q6HES7	Q6hes7 bacillus th
316	22	91.7	334	2	O50178	O50178 pseudomonas	389	22	91.7	365	2	Q73212	Q73212 bacillus ce
317	22	91.7	334	2	Q885J1	Q885j1 pseudomonas	390	22	91.7	365	2	Q81WF1	Q81wf1 bacillus an
318	22	91.7	334	2	Q9RT96	Q9rt96 deinococcus	391	22	91.7	365	2	Q71BH6	Q71bh6 human papil
319	22	91.7	334	2	Q8YUB2	Q8yub2 anabaena sp	392	22	91.7	365	2	Q71Bi4	Q71bi4 human papil
320	22	91.7	334	2	Q7DCH1	Q7dch1 pseudomonas	393	22	91.7	365	2	Q71c70	Q71c70 human papil
321	22	91.7	335	1	SYFA_PROMP	Q7v0i7 prochloroco	394	22	91.7	365	2	Q8B5P3	Q8b5p3 human papil
322	22	91.7	335	2	Q6LUI5	Q6lu15 photobacter	395	22	91.7	365	2	Q8JQA8	Q8jq8a human papil
323	22	91.7	335	2	CAG19040	Cag19040 photobact	396	22	91.7	365	2	Q9WQX4	Q9wqx4 human papil

337	22	91.7	365	2	Q9YIV0	Q9Yiv0 human papil	470	22	91.7	401	2	Q9U276	Q9u276 caenorhabdi
338	22	91.7	365	2	Q9YV72	Q9Yv72 human papil	471	22	91.7	401	2	Q8KIM2	Q8klm2 rattus norv
339	22	91.7	365	2	Q9YV73	Q9Yv73 human papil	472	22	91.7	402	2	Q6SGX7	Q6sgx7 uncultured
400	22	91.7	365	2	Q9YV74	Q9Yv74 human papil	473	22	91.7	402	2	Q8KIM3	Q8klm3 mus musculu
401	22	91.7	365	2	Q9YV75	Q9Yv75 human papil	474	22	91.7	402	2	AAR37842	Aar37842 unculture
402	22	91.7	365	2	Q9YV76	Q9Yv76 human papil	475	22	91.7	403	1	KAP2_HUMAN	Pl3861 homo sapien
403	22	91.7	365	2	Q9E839	Q9E839 human papil	476	22	91.7	403	2	Q8NKM8	Q8nkm8 uncultured
404	22	91.7	365	2	Q9E840	Q9E840 human papil	477	22	91.7	405	1	I1BC_HORSE	Q9cv13 equus cabal
405	22	91.7	365	2	Q9E841	Q9E841 human papil	478	22	91.7	405	2	Q9XB04	Q9xbu4 bacillus ce
406	22	91.7	365	2	Q9E842	Q9E842 human papil	479	22	91.7	405	2	Q8VM49	Q8vm49 bacillus th
407	22	91.7	365	2	Q9E842	Q9E842 human pap	480	22	91.7	405	2	O07191	O07191 mycobacteri
408	22	91.7	365	2	AAQ03824	AAQ03824 human pap	481	22	91.7	405	2	Q7TY38	Q7ty38 mycobacteri
409	22	91.7	365	2	AAQ10406	AAQ10406 human pap	482	22	91.7	408	2	Q9VLI0	Q9vli0 drosophila
410	22	91.7	365	2	AAQ10715	AAQ10715 human pap	483	22	91.7	410	2	Q9M379	Q9m379 arabidopsis
411	22	91.7	365	2	AAQ10723	AAQ10723 human pap	484	22	91.7	410	2	Q93X40	Q93x40 lactuca sat
412	22	91.7	365	2	AAQ42835	AAQ42835 bacillus	485	22	91.7	410	2	Q9RPE5	Q9rfes rhodobacter
413	22	91.7	365	2	AAQ33143	AAQ33143 bacillus	486	22	91.7	410	2	Q737F0	Q737f0 bacillus ce
414	22	91.7	366	1	MINC_ANASP	Q9Yrj1 anabaena sp	487	22	91.7	410	2	AAQ41612	AAQ41612 bacillus
415	22	91.7	366	2	Q7MI89	Q7mi89 vibrio vuln	488	22	91.7	411	2	Q94252	Q94252 caenorhabdi
416	22	91.7	366	2	Q8DBN0	Q8dbn0 vibrio vuln	489	22	91.7	413	1	NEUA_STRA3	Q53598 streptococc
417	22	91.7	367	2	Q7Q1Q7	Q7q1q7 anopheles g	490	22	91.7	413	1	NEUA_STRA5	Q9afg9 streptococc
418	22	91.7	368	2	Q93WB9	Q93wb9 oryza sativ	491	22	91.7	413	2	Q9ALM4	Q9alw4 streptococc
419	22	91.7	370	1	SERC_METAC	Q8nnil methanosarc	492	22	91.7	413	2	Q93TI0	Q93ti0 streptococc
420	22	91.7	370	1	YB11_HALNL	Pl7103 halobacteri	493	22	91.7	415	1	LEU2_PYRAE	Q8xw0 pyrobaculum
421	22	91.7	370	2	Q9VG48	Q9vg48 drosophila	494	22	91.7	419	2	Q9XVE0	Q9xve0 caenorhabdi
422	22	91.7	372	1	VE2_HPV31	Pl7383 human papil	495	22	91.7	420	2	Q7S229	Q7s229 neurospora
423	22	91.7	373	1	Y007_METUA	Q60318 methanococc	496	22	91.7	420	2	O87350	O87350 lactococcus
424	22	91.7	373	2	Q9NEU3	Q9neu3 caenorhabdi	497	22	91.7	420	2	Q7BN83	Q7bn83 lactococcus
425	22	91.7	373	2	Q819S2	Q819s2 bacillus ce	498	22	91.7	421	2	Q9ACM3	Q9acm3 streptococc
426	22	91.7	373	2	Q69342	Q69342 suid herpes	499	22	91.7	423	2	Q6VMI7	Q6vml7 streptomyce
427	22	91.7	375	2	Q93XV6	Q9qgm8 suid herpes	500	22	91.7	423	2	Q7P0B7	Q7p0b7 chromobacte
428	22	91.7	376	1	C1SY_PYRFU	Q53554 pyrococcus	501	22	91.7	423	2	Q8G6S4	Q8g6s4 bifidobacte
429	22	91.7	376	1	YOGT_BACSU	P54497 bacillus su	502	22	91.7	423	2	AAR30152	Aar30152 streptomy
430	22	91.7	377	1	ICED_BOVIN	Q75601 bos taurus	503	22	91.7	424	2	Q84CJ7	Q84cj7 streptomyce
431	22	91.7	378	2	Q9FXW6	Q9fxw6 arabidopsis	504	22	91.7	431	1	HISX_XANAC	Q8plg9 xanthomonas
432	22	91.7	378	2	Q89VE1	Q89ve1 bradyrhizob	505	22	91.7	431	1	HISX_XANCP	Q8p9p2 xanthomonas
433	22	91.7	379	2	Q9VEA6	Q9vea6 drosophila	506	22	91.7	431	2	Q7WJQ2	Q7wlq2 bordetella
434	22	91.7	379	2	Q9LXE2	Q9lxe2 arabidopsis	507	22	91.7	432	2	Q18418	Q18418 caenorhabdi
435	22	91.7	381	1	PRC1_CORGL	Q8nsh7 corynebacte	508	22	91.7	432	2	Q924X3	Q924x3 streptomyce
436	22	91.7	382	2	Q9BUB1	Q9bub1 homo sapien	509	22	91.7	434	2	Q9HJ87	Q9hj87 thermoplasm
437	22	91.7	382	2	Q84HN9	Q84hn9 streptomyce	510	22	91.7	434	2	Q7D6B9	Q7d6b9 mycobacteri
438	22	91.7	383	1	PRC2_CORGL	Q8nsl1 corynebacte	511	22	91.7	434	2	Q6HI57	Q6hi57 bacillus th
439	22	91.7	384	1	C1SY_THEAC	P21553 thermoplasm	512	22	91.7	434	2	Q81PX7	Q81px7 bacillus an
440	22	91.7	384	2	Q6EGQ6	Q6egq6 human papil	513	22	91.7	436	2	Q8Q0L0	Aat31779 bacillus
441	22	91.7	385	2	Q8TZ62	Q8tz62 methanopyru	514	22	91.7	436	2	Q8TKV7	Q8tkv7 methanosarc
442	22	91.7	385	2	Q94CF8	Q94cf8 capsicum an	515	22	91.7	439	2	Q9A008	Q9a0q8 streptococc
443	22	91.7	386	2	Q97C64	Q97c64 thermoplasm	516	22	91.7	439	2	Q9A008	Q9a0q8 streptococc
444	22	91.7	388	2	Q86P44	Q86p44 drosophila	517	22	91.7	442	2	Q7XV19	Q7xv19 oryza sativ
445	22	91.7	389	1	ALR1_BACSU	P10725 bacillus su	518	22	91.7	447	2	Q6J5M9	Q6j5m9 solanum tub
446	22	91.7	389	2	Q976X3	Q976x3 sulfolobus	519	22	91.7	447	2	Q8W4B0	Q8w4b0 arabidopsis
447	22	91.7	389	2	Q97U78	Q97u78 sulfolobus	520	22	91.7	447	2	Q8AB03	Q8ab03 bacteroides
448	22	91.7	389	2	Q72W91	Q72w91 leptospira	521	22	91.7	447	2	AAT37527	Aat37527 solanum t
449	22	91.7	389	2	Q8FY95	Q8fy95 leptospira	522	22	91.7	452	2	Q9YFU6	Q9yfu6 aeropyrum p
450	22	91.7	389	2	AA568683	AA568683 leptospir	523	22	91.7	452	2	Q7XPS8	Q7xps8 oryza sativ
451	22	91.7	390	2	Q20910	Q20910 caenorhabdi	524	22	91.7	454	2	Q7XPS8	Q7xps8 fusarium so
452	22	91.7	390	2	Q72TP8	Q72tp8 leptospira	525	22	91.7	457	1	CGPB_FUSSO	Q00858 fusarium so
453	22	91.7	390	2	Q8F1J1	Q8f1j1 leptospira	526	22	91.7	458	2	Q7QKW1	Q7qkw1 anopheles g
454	22	91.7	390	2	AA569580	AA569580 leptospir	527	22	91.7	460	1	LEU2_LACLA	Q02142 lactococcus
455	22	91.7	391	2	Q84B24	Q84b24 thermus the	528	22	91.7	460	2	Q93VM7	Q93vm7 lupinus alb
456	22	91.7	391	2	Q72K30	Q72k30 thermus the	529	22	91.7	460	2	Q7USR0	Q7u5r0 synecchococc
457	22	91.7	391	2	AA580936	AA580936 thermus t	530	22	91.7	461	2	Q9S9N6	Q9s9n6 arabidopsis
458	22	91.7	393	1	Q8DTM1	Q8dtm1 streptococc	531	22	91.7	462	1	LEU2_LISIN	Q92a26 listeria in
459	22	91.7	394	1	MPGS_PYRAB	Q9uzcl pyrococcus	532	22	91.7	462	1	LEU2_LISMO	Q9y5r7 listeria mo
460	22	91.7	396	1	YB0F_SCHPO	P87156 schizosacch	533	22	91.7	462	2	Q9LRE7	Q9lre7 lupinus alb
461	22	91.7	398	1	ACK1_VIBCH	Q9kt07 vibrio chol	534	22	91.7	462	2	Q71V33	Q71v33 listeria mo
462	22	91.7	398	1	ACK1_VIBFA	Q8mz4 vibrio para	535	22	91.7	463	2	AA04782	Aat04782 listeria
463	22	91.7	398	1	ACK1_VIBVU	Q8dah8 vibrio vuln	536	22	91.7	463	2	Q8L6L1	Q8l6l1 lupinus lut
464	22	91.7	398	2	Q7MJM8	Q7mjm8 vibrio vuln	537	22	91.7	463	2	Q7ZYE4	Q7zyb4 xenopus lae
465	22	91.7	398	2	Q98PF1	Q98pf1 mycoplasma	538	22	91.7	465	2	Q9SDZ9	Q9sdz9 ipomoea bat
466	22	91.7	398	2	Q6LNF6	Q6lnf6 photobacter	539	22	91.7	465	2	Q9ZSS0	Q9zss0 ipomoea bat
467	22	91.7	398	2	CAG21170	Cag21170 photobact	540	22	91.7	466	1	SELA_RHIME	P58226 rhizobium m
468	22	91.7	400	1	KAP2_BOVIN	P05115 bos taurus	541	22	91.7	466	2	Q9LRJ3	Q9lrj3 tagetes pat
469	22	91.7	400	1	KAP2_MOUSE	P12367 mus musculu	542	22	91.7	466	2	Q9CH26	Q9ch26 lactococcus

543	22	91.7	468	2	Q84KZ1	Q84kz1 nicotiana t	616	22	91.7	546	2	O6GCT8	Q6act8 staphylococ
544	22	91.7	468	2	Q8SIV9	Q8siv9 arabidopsis	617	22	91.7	546	2	O6GK38	Q6skb8 staphylococ
545	22	91.7	469	2	Q723Z8	Q723z8 listeria mo	618	22	91.7	546	2	Q8NYM2	Q8nym2 staphylococ
546	22	91.7	469	2	Q8K03113	Aat03113 listeria	619	22	91.7	546	2	Q9X333	Q99x33 staphylococ
547	22	91.7	470	2	Q84KZ2	Q84kz2 nicotiana t	620	22	91.7	546	2	Q7A808	Q7a808 staphylococ
548	22	91.7	470	2	Q9XF09	Q9xf09 anchusa off	621	22	91.7	547	2	Q6QBS4	Q6qbs4 lactococcus
549	22	91.7	470	2	Q8EIV4	Q8eiv4 shewanella	622	22	91.7	547	2	AA49166	AA49166 lactococc
550	22	91.7	471	2	Q8AZJ4	Q8azj4 human herpe	623	22	91.7	549	2	Q8CQ11	Q8cgy1 staphylococ
551	22	91.7	472	1	YAHG_ECOLI	P77221 escherichia	624	22	91.7	550	2	P71323	P71323 pantoea agg
552	22	91.7	472	2	Q6CE61	Q6ce61 yarrowia li	625	22	91.7	550	2	Q9FDC2	Q9fdc2 pseudomonas
553	22	91.7	474	1	LEU2_XYLFA	Q9pax0 xyella fas	626	22	91.7	550	2	Q8Z4X7	Q8z4x7 salmonella
554	22	91.7	474	1	LEU2_XYLFT	Q87bp9 xyella fas	627	22	91.7	550	2	Q93IM7	Q93im7 salmonella
555	22	91.7	474	2	Q7NI60	Q7ni60 gloeobacter	628	22	91.7	550	2	Q7CQ30	Q7cq30 salmonella
556	22	91.7	474	2	Q92092	Q92092 rat cytomeg	629	22	91.7	552	1	DCIP_ENTCL	P22334 enterobacte
557	22	91.7	475	1	SW_BOVIN	P17248 bos taurus	630	22	91.7	552	2	Q6MJM2	Q6mjm2 bdellovibri
558	22	91.7	476	2	Q7NW21	Q7nw21 chromobacte	631	22	91.7	552	2	CAE80538	CAE80538 bdellovib
559	22	91.7	477	1	LE21_BRAJA	Q89x98 bradyrhizob	632	22	91.7	553	1	HUTU_STAAM	Q931g1 staphylococ
560	22	91.7	477	1	PBP4_ECOLI	P24228 escherichia	633	22	91.7	553	1	HUTU_STAAN	Q99ru2 staphylococ
561	22	91.7	477	2	Q7AA10	Q7aa10 escherichia	634	22	91.7	553	1	VH65_NPVAC	Q85339 autographa
562	22	91.7	477	2	Q8CPV2	Q8cpv2 escherichia	635	22	91.7	553	1	Q87019	Q870y9 neurospora
563	22	91.7	477	2	Q8X9K8	Q8x9k8 escherichia	636	22	91.7	553	2	Q6CCR8	Q6ccr8 yarrowia li
564	22	91.7	480	2	Q7V8L8	Q7v8l8 prochloroc	637	22	91.7	553	2	Q9FVY9	Q9fvy9 oryza sativ
565	22	91.7	481	2	Q76YW5	Q76yw5 bacterioph	638	22	91.7	553	2	Q6G6Y9	Q6g6y9 staphylococ
566	22	91.7	481	2	AAQ17781	AAq17781 bacteriop	639	22	91.7	553	2	Q6GEA4	Q6gea4 staphylococ
567	22	91.7	486	2	Q8FRY2	Q8fry2 corynebacte	640	22	91.7	554	1	HUTU_THEAC	Q9hli9 thermoplas
568	22	91.7	487	2	Q9U3A4	Q9u3a4 caenorhabdi	641	22	91.7	555	2	Q6P9H4	Q6p9h4 homo sapien
569	22	91.7	488	2	Q9LZY4	Q9lzy4 arabidopsis	642	22	91.7	555	2	Q9FN65	Q9fn65 homo sapien
570	22	91.7	489	2	Q7WBMS	Q7wbms bordetella	643	22	91.7	555	2	Q6D143	Q6d143 erwinia car
571	22	91.7	489	2	Q7WN49	Q7wn49 bordetella	644	22	91.7	555	2	Q8BMA3	Q8bma3 m mus muscu
572	22	91.7	492	2	Q8CXF6	Q8cx6 oceanobacil	645	22	91.7	555	2	AAH60761	AAh60761 homo sapi
573	22	91.7	493	2	Q84609	Q84609 arabidopsis	646	22	91.7	557	1	SYQ_HAEIN	P43831 haemophilus
574	22	91.7	493	2	Q9LRA5	Q9lra5 arabidopsis	647	22	91.7	558	2	Q81DD4	Q81dd4 bacillus ce
575	22	91.7	494	2	Q8ENM2	Q8enm2 oceanobacil	648	22	91.7	560	1	TDRH_MOUSE	Q80v11 mus musculus
576	22	91.7	496	2	Q8LL67	Q8ll67 amaranthus	649	22	91.7	560	2	Q7D958	Q7d958 mycobacteri
577	22	91.7	500	2	Q97E43	Q97e43 clostridium	650	22	91.7	560	2	O53865	O53865 mycobacteri
578	22	91.7	501	2	O57693	O57693 thermoprote	651	22	91.7	560	2	Q7U140	Q7u140 mycobacteri
579	22	91.7	501	2	Q6GMJ2	Q6gmj2 brachydanio	652	22	91.7	560	2	Q6FGS8	Q6fpg8 mus musculus
580	22	91.7	504	1	GALI_CANPA	O42821 candida par	653	22	91.7	560	2	AAH57030	AAh57030 mus muscu
581	22	91.7	508	2	Q6YMQ3	Q6ymq3 oryza sativ	654	22	91.7	561	2	Q6HIMI	Q6him1 bacillus th
582	22	91.7	508	2	BAD15926	Bad15926 oryza sat	655	22	91.7	561	2	Q737X8	Q737x8 bacillus ce
583	22	91.7	508	2	BAD16365	Bad16365 oryza sat	656	22	91.7	561	2	Q91Q80	Q91qe0 bacillus an
584	22	91.7	510	2	Q8SQ07	Q8sqg7 arabidopsis	657	22	91.7	561	2	AA41434	AA41434 bacillus
585	22	91.7	516	2	Q6NW94	Q6nw94 brachydanio	658	22	91.7	561	2	AAAT31600	AAat31600 bacillus
586	22	91.7	516	2	Q6PHF8	Q6phf8 brachydanio	659	22	91.7	563	2	Q742Q2	Q742q2 mycobacteri
587	22	91.7	516	2	AAH56567	AAh56567 brachydan	660	22	91.7	563	2	AA5031100	AA503100 mycobacte
588	22	91.7	516	2	AAH67675	AAh67675 brachydan	661	22	91.7	565	2	Q6M6B5	Q6m6b5 corynebacte
589	22	91.7	517	2	Q8P8Y3	Q8p8y3 xanthomonas	662	22	91.7	565	2	CAP19755	Cap19755 corynebac
590	22	91.7	520	2	Q72SW5	Q72sw5 leptospira	663	22	91.7	568	1	HRPA_RALSO	O52498 ralstonia s
591	22	91.7	520	2	Q8F2L0	Q8f2l0 leptospira	664	22	91.7	571	2	Q51763	Q51763 pseudomonas
592	22	91.7	520	2	AA569863	AA569863 leptospir	665	22	91.7	572	2	Q8B9F4	Q8b9f4 rachiplusia
593	22	91.7	521	2	Q98JA5	Q98ja5 rhizobium l	666	22	91.7	575	2	Q84Q92	Q84q92 oryza sativ
594	22	91.7	523	2	Q8PY13	Q8py13 methanosarc	667	22	91.7	575	2	Q8H405	Q8h405 oryza sativ
595	22	91.7	523	2	Q8THZ6	Q8thz6 methanosarc	668	22	91.7	576	2	Q7QH44	Q7qh44 anopheles g
596	22	91.7	525	2	Q7CV01	Q7cv01 agrobacteri	669	22	91.7	577	2	Q6XAE8	Q6xae8 saccharomyc
597	22	91.7	525	2	Q8U7J6	Q8u7j6 agrobacteri	670	22	91.7	577	2	Q8BJR2	Q8bjr2 mus musculus
598	22	91.7	528	2	Q8SVJ8	Q8svj8 encephalito	671	22	91.7	577	2	Q8BJT9	Q8bjt9 mus musculus
599	22	91.7	530	1	UDBH_HUMAN	Q75795 homo sapien	672	22	91.7	577	2	Q8BK85	Q8bk85 mus musculus
600	22	91.7	532	2	Q84148	Q84148 buchnera ap	673	22	91.7	577	2	Q91VV3	Q91vv3 mus musculus
601	22	91.7	532	2	Q84149	Q84149 buchnera ap	674	22	91.7	577	2	AAp75900	AAp75900 saccharom
602	22	91.7	532	2	Q84150	Q84i50 buchnera ap	675	22	91.7	578	1	CT31_HUMAN	Q9bv94 homo sapien
603	22	91.7	533	2	Q86173	Q86i73 mycobacteri	676	22	91.7	578	2	AAH16184	AAh16184 homo sapi
604	22	91.7	533	2	Q7TYF0	Q7tyf0 mycobacteri	677	22	91.7	578	2	CAG33547	Cag33547 homo sapi
605	22	91.7	534	2	Q8ZUW2	Q8zuw2 pyrobaculum	678	22	91.7	579	2	Q8NRJ7	Q8nrj7 corynebacte
606	22	91.7	534	2	Q76YL0	Q76yl0 bacterioph	679	22	91.7	585	2	Q7QBT2	Q7qbt2 anopheles g
607	22	91.7	534	2	Q6D0E8	Q6d0e8 erwinia car	680	22	91.7	588	2	Q8TPV4	Q8tpv4 methanosarc
608	22	91.7	534	2	AA61695	AA61695 bacteriop	681	22	91.7	590	2	Q8SSB8	Q8ssb8 encephalito
609	22	91.7	537	2	Q7QPS7	Q7qps7 giardia lam	682	22	91.7	594	2	Q87ZJ7	Q87zj7 pseudomonas
610	22	91.7	538	1	PYRG_AERPE	Q9ybj4 aeropyrum p	683	22	91.7	602	2	Q6VP52	Q6vp52 candidatu
611	22	91.7	541	2	Q6UWZ4	Q6uwz4 homo sapien	684	22	91.7	602	2	AA99265	AA99265 candidatu
612	22	91.7	541	2	AAQ88943	AAc88943 homo sapi	685	22	91.7	607	2	Q8Q077	Q8q077 methanosarc
613	22	91.7	543	1	ARD1_CAEL	Q95654 caenorhabdi	686	22	91.7	607	2	Q9LFA7	Q9lfa7 arabidopsis
614	22	91.7	543	2	Q9SKC3	Q9skc3 arabidopsis	687	22	91.7	608	2	Q6CEX5	Q6cex5 yarrowia li
615	22	91.7	545	2	Q8S9K1	Q8s9k1 arabidopsis	688	22	91.7	615	2	Q98LK3	Q98lk3 rhizobium l



835	22	91.7	901	2	Q8A5P7	Q8A5p7 bacteroides	908	22	91.7	1248	2	Q6N9K1	Q6n9k1 rhodopseudo
836	22	91.7	901	2	Q9DXA2	Q9dxax2 choristoneu	909	22	91.7	1248	2	CAE26985	CAe26985 rhodopseu
837	22	91.7	902	2	Q6C3A4	Q6c3a4 yarrowia li	910	22	91.7	1270	1	MYPC MOUSE	O70468 mus musculus
838	22	91.7	904	2	Q8A1K1	Q8a1k1 bacteroides	911	22	91.7	1273	2	Q6R2F7	Q6r2f7 homo sapien
839	22	91.7	910	2	Q96EK7	Q96ek7 homo sapien	912	22	91.7	1273	2	AAR89909	Aar89909 homo sapi
840	22	91.7	914	2	Q91W15	Q91w15 mus musculus	913	22	91.7	1274	1	MYPC HUMAN	Q14896 homo sapien
841	22	91.7	917	2	Q96J19	Q96j19 homo sapien	914	22	91.7	1274	2	Q9UM53	Q9um53 homo sapien
842	22	91.7	920	2	Q28984	Q28984 sus scrofa	915	22	91.7	1280	2	Q22554	Q22554 caenorhabdi
843	22	91.7	924	2	Q6MJP4	Q6mjp4 bdellovibri	916	22	91.7	1282	2	Q46348	Q46348 cytophaga s
844	22	91.7	924	2	CAE80516	CAe80516 bdellovib	917	22	91.7	1282	2	Q90X86	Q90x86 xenopus lae
845	22	91.7	928	2	Q6N3P6	Q6n3p6 rhodopseu	918	22	91.7	1289	1	DUET HUMAN	Q9y2a5 homo sapien
846	22	91.7	928	2	CAE29088	CAe29088 rhodopseu	919	22	91.7	1290	2	Q8Y5J3	Q8y5j3 anabaena sp
847	22	91.7	930	1	FRSK_PASMU	Q9cpl3 pasteurella	920	22	91.7	1292	2	Q9ZG85	Q9zge5 heliobacilli
848	22	91.7	947	2	QCS6F9	Q786f9 neurospora	921	22	91.7	1294	2	Q7RNG8	Q7rnga8 plasmodium
849	22	91.7	954	1	GSCP RHIME	Q92q11 rhizobium m	922	22	91.7	1339	1	DPOA_TRYBB	P27727 trypanosoma
850	22	91.7	964	1	RPO MCMV	P11640 maize chlor	923	22	91.7	1342	2	Q9VPZ7	Q9vpz7 drosophila
851	22	91.7	965	2	Q91BM2	Q91bm2 maize chlor	924	22	91.7	1342	2	Q9GPP6	Q9gpp6 drosophila
852	22	91.7	974	2	Q8S905	Q8s905 arabidopsis	925	22	91.7	1352	2	Q9ZUF0	Q9zuf0 arabidopsis
853	22	91.7	988	2	Q7XAC4	Q7xac4 arabidopsis	926	22	91.7	1374	2	Q22148	Q22148 arabidopsis
854	22	91.7	988	2	Q7PC79	Q7pc79 arabidopsis	927	22	91.7	1395	2	Q44924	Q44924 drosophila
855	22	91.7	993	2	Q9CAH2	Q9cah2 arabidopsis	928	22	91.7	1395	2	Q7KVK3	Q7kvk3 drosophila
856	22	91.7	1003	2	Q9LPQ5	Q9lpq5 arabidopsis	929	22	91.7	1395	2	AAF46887	Aaf46887 drosophil
857	22	91.7	1012	2	Q7S3H9	Q7s3h9 neurospora	930	22	91.7	1412	2	O64612	O64612 arabidopsis
858	22	91.7	1019	2	Q7TN13	Q7tn13 mus musculus	931	22	91.7	1427	2	Q9VZT8	Q9vzt8 drosophila
859	22	91.7	1021	2	Q9RUL0	Q9rul0 deinococcus	932	22	91.7	1429	2	Q9WZ13	Q9wzt13 drosophila
860	22	91.7	1031	2	Q7RON3	Q7ron3 giardia lam	933	22	91.7	1429	2	AAW71113	AAw71113 drosophil
861	22	91.7	1035	2	Q6NBZ7	Q6nbz7 rhodopseu	934	22	91.7	1442	2	Q6P9L1	Q6p9l1 mus muscu
862	22	91.7	1035	2	CAE26122	CAe26122 rhodopseu	935	22	91.7	1442	2	AAH60717	Aah60717 mus muscu
863	22	91.7	1036	2	O64819	O64819 arabidopsis	936	22	91.7	1504	2	O6RCR0	O6rcr0 legionella
864	22	91.7	1040	1	I2 BACTN	Q8a2a1 bacteroides	937	22	91.7	1504	2	AAR90356	Aar90356 legionell
865	22	91.7	1043	2	Q97E41	Q97e41 clostridium	938	22	91.7	1513	2	Q76DQ2	Q76dq2 legionella
866	22	91.7	1047	2	Q9A821	Q9a821 caulobacter	939	22	91.7	1513	2	BAD06870	Bad06870 legionell
867	22	91.7	1050	2	Q97X88	Q97xx8 sulfobus	940	22	91.7	1538	2	Q6RCQ8	Q6rcq8 legionella
868	22	91.7	1050	2	Q8PU41	Q8pu41 methanosarc	941	22	91.7	1538	2	AAR90358	Aar90358 legionell
869	22	91.7	1052	2	O6BS66	O6bs66 debaryomyce	942	22	91.7	1556	2	Q7Y1P7	Q7y1p7 oryza sativ
870	22	91.7	1061	2	Q6MQS6	Q6mgs6 bdellovibri	943	22	91.7	1566	2	Q7NEB4	Q7neb4 gloebacter
871	22	91.7	1061	2	CAE78371	CAe78371 bdellovib	944	22	91.7	1590	1	GCN2_YEARST	P15442 saccharomyc
872	22	91.7	1063	2	Q7PMK4	Q7pmk4 anopheles g	945	22	91.7	1630	2	Q6PCS0	Q6pcso mus musculus
873	22	91.7	1064	1	SVI AERPE	Q9yvf67 aeropyrum p	946	22	91.7	1630	2	AAH59192	Aah59192 mus muscu
874	22	91.7	1071	1	CARB BACSU	P25994 bacillus su	947	22	91.7	1635	2	Q88L26	Q88l26 pseudomonas
875	22	91.7	1072	2	Q9LVQ1	Q9lvq1 arabidopsis	948	22	91.7	1652	2	Q7XE51	Q7xe51 oryza sativ
876	22	91.7	1075	2	Q9H2G1	Q9h2g1 homo sapien	949	22	91.7	1662	2	Q8T6J2	Q8t6j2 dictyosteli
877	22	91.7	1081	1	IP04 HUMAN	Q8tex9 homo sapien	950	22	91.7	1787	2	O26810	O26810 methanobact
878	22	91.7	1082	1	IP04 MOUSE	Q8vi75 mus musculus	951	22	91.7	1807	1	ITB4 RAT	Q64632 rattus norv
879	22	91.7	1091	2	Q8Y4J2	Q8y4j2 listeria mo	952	22	91.7	1822	2	ITB4 HUMAN	P16144 homo sapien
880	22	91.7	1093	2	Q7P5E9	Q7p5e9 fusobacteri	953	22	91.7	1868	2	Q6H242	Q6h242 lactobacilli
881	22	91.7	1102	2	Q6LP30	Q6lp30 xenopus lae	954	22	91.7	1908	2	Q7MT05	Q7mt05 porphyronon
882	22	91.7	1102	2	AAR72089	Aar72089 xenopus l	955	22	91.7	1926	2	Q6RCQ9	Q6rcq9 legionella
883	22	91.7	1111	2	Q73I04	Q73i04 wolbachia p	956	22	91.7	1926	2	AAR90357	Aar90357 legionell
884	22	91.7	1111	2	AA514108	AAa14108 wolbachia	957	22	91.7	1933	2	Q9YEF6	Q9yef6 aeropyrum p
885	22	91.7	1113	2	Q75D66	Q75d66 ashbya goss	958	22	91.7	1978	2	Q14966	Q14966 homo sapien
886	22	91.7	1113	2	AA550929	AAa50929 ashbya go	959	22	91.7	1978	2	Q7Z3T7	Q7z3t7 homo sapien
887	22	91.7	1116	2	Q7CW80	Q7cw80 agrobacteri	960	22	91.7	2024	2	Q7PZ21	Q7pzz1 anopheles g
888	22	91.7	1116	2	Q9UBP7	Q9ubp7 agrobacteri	961	22	91.7	2152	2	Q9SND1	Q9snd1 arabidopsis
889	22	91.7	1117	2	Q9CLG5	Q6clg5 kluyveromyc	962	22	91.7	2157	2	Q8LHF0	Q8lhf0 oryza sativ
890	22	91.7	1117	2	Q9U9K7	Q9u9k7 caenorhabdi	963	22	91.7	2166	2	Q9ZNX7	Q9znx7 oryza sativ
891	22	91.7	1121	2	Q8W302	Q8w302 oryza sativ	964	22	91.7	2194	1	GLSN MEDSA	Q03460 medicago sa
892	22	91.7	1132	2	Q9W475	Q9w475 drosophila	965	22	91.7	2194	2	O40360	Q40360 medicago gla
893	22	91.7	1133	2	Q8RDW1	Q8rdw1 fusobacteri	966	22	91.7	2211	2	O6FLJ2	O6flj2 candida gla
894	22	91.7	1139	2	Q8NF92	Q8nf92 homo sapien	967	22	91.7	2211	2	Q8BYD4	Q8byd4 debaryomyce
895	22	91.7	1148	2	Q6CBJ2	Q6cbj2 yarrowia li	968	22	91.7	2216	2	Q9LV03	Q9lv03 arabidopsis
896	22	91.7	1154	2	Q7IG06	Q810g6 caenorhabdi	969	22	91.7	2225	1	PYR1 MESAU	P08955 mesocricetu
897	22	91.7	1165	2	Q72TU4	Q72tu4 leptospira	970	22	91.7	2241	2	Q7XRJ2	Q7xrj2 oryza sativ
898	22	91.7	1165	2	Q8FLC9	Q8flc9 leptospira	971	22	91.7	2258	2	Q7QF47	Q7qf47 anopheles g
899	22	91.7	1165	2	AA569534	AAa69534 leptospir	972	22	91.7	2307	2	Q9AG79	Q9ag79 streptomyc
900	22	91.7	1175	2	Q9BLU7	Q9blu7 rhizobium l	973	22	91.7	2359	2	Q9JIF1	Q9jif1 rattus norv
901	22	91.7	1181	2	O17646	O17646 caenorhabdi	974	22	91.7	3597	2	Q8ILR5	Q8ilr5 plasmodium
902	22	91.7	1190	2	Q9FNB4	Q9fnb4 arabidopsis	975	22	91.7	3906	2	Q8G987	Q8g987 oscillatori
903	22	91.7	1193	2	Q93W04	Q93w04 oryza sativ	976	22	91.7	3906	2	Q9RNB2	Q9rnb2 microcystis
904	22	91.7	1200	2	Q9WXB6	Q9wxb6 acidiphiliu	977	22	91.7	3906	2	O9FDU1	O9fdul microcystis
905	22	91.7	1229	2	Q8ZY40	Q8zy40 nitrosomona	978	22	91.7	5636	2	Q9N9M2	Q9nm2 leishmania
906	22	91.7	1236	2	Q9UPA4	Q9upa4 rhodocyclus	979	22	91.7	18412	2	Q7ZZ61	Q7zz61 brachydanio
907	22	91.7	1237	2	Q91976	Q91976 gallus gall	980	22	91.7	19066	2	Q801W8	Q801w8 brachydanio

981 22 91.7 26926 2 Q10466  
 982 22 91.7 26926 2 Q8WZB3  
 983 22 91.7 34350 2 Q8WZ42  
 984 21 87.5 15 1 RM12\_YEAST  
 985 21 87.5 23 2 Q9UMJ6  
 986 21 87.5 27 2 Q79441  
 987 21 87.5 27 2 Q79859  
 988 21 87.5 40 2 Q7VLC7  
 989 21 87.5 51 2 Q39569  
 990 21 87.5 51 2 Q84LA3  
 991 21 87.5 51 2 Q9SCW2  
 992 21 87.5 51 2 Q96526  
 993 21 87.5 53 2 Q6RU04  
 994 21 87.5 53 2 Q6RU06  
 995 21 87.5 53 2 AAR84407  
 996 21 87.5 53 2 AAR84409  
 997 21 87.5 54 2 Q24519  
 998 21 87.5 54 2 Q43625  
 999 21 87.5 56 2 Q9AZH6  
 1000 21 87.5 57 1 YG21\_METJA

## ALIGNMENTS

## RESULT 1

Q8W6Q0 PRELIMINARY; PRT; 47 AA.  
 AC Q8W6Q0;  
 DT 01-WAR-2002 (TRENBLrel. 20, Created)  
 DT 01-WAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-WAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Gp50.  
 OS Bacteriophage phiE125.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
 OC Lambda-like viruses.  
 OC NCBI\_TaxID=180504;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=22077271; PubMed=12081973;  
 RA Woods D.E., Jeddeloh J.A., Fritz D.L., DeShazer D.;  
 RT "Burkholderia thailandensis E125 harbors a temperate bacteriophage  
 RT specific for Burkholderia mallei.";  
 RL J. Bacteriol. 184:4003-4017(2002).  
 DR EMBL; AF447491; AAL40324.1; -  
 SQ SEQUENCE 47 AA; 5016 MW; A9E8258058E9F5DF CRC64;

Query Match 91.7%; Score 22; DB 2; Length 47;  
 Best Local Similarity 80.0%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

Db 10 GTYVE 14

## RESULT 2

Q8LLV4 PRELIMINARY; PRT; 51 AA.  
 ID Q8LLV4  
 AC Q8LLV4;  
 DT 01-OCT-2002 (TRENBLrel. 22, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
 DE UORF.  
 GN Name=SAMDC;  
 OS Ipomoea batatas (Sweet potato) (Batate).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Convolvulaceae; Ipomoeae; Ipomoea.  
 OC NCBI\_TaxID=4120;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Leaf;

RA Chiang W.J., Chen S.C.G.;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF291761; AAN03493.1; -  
 SQ SEQUENCE 51 AA; 5544 MW; PDEF59932C05AFCA CRC64;

Query Match 91.7%; Score 22; DB 2; Length 51;  
 Best Local Similarity 80.0%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

Db 23 GTYVE 27

## RESULT 3

Q8U8S6 PRELIMINARY; PRT; 76 AA.  
 ID Q8U8S6;  
 AC Q8U8S6;  
 DT 01-JUN-2002 (TRENBLrel. 21, Created)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Transcriptional regulator.  
 GN OrderedLocusNames=Atu4014;  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 OC NCBI\_TaxID=176299;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Dupont;  
 RC MEDLINE=21608550; PubMed=11743193;  
 RX Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RA "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 RT C58.";  
 RT Science 294:2317-2323(2001).  
 RL EMBL; AB009332; AAL44815.1; -  
 DR PIR; A13049; A13049.  
 DR PIR; B98236; B98236.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR InterPro; IPR01387; HTH\_3.  
 DR InterPro; IPR010982; Lambda\_like\_DNA.  
 DR SMART; SM00530; HTH\_XRE; 1.  
 KW Complete proteome.

SQ SEQUENCE 76 AA; 8222 MW; BD4C0D17101E94AA CRC64;

Query Match 91.7%; Score 22; DB 2; Length 76;  
 Best Local Similarity 80.0%; Pred. No. 8.7e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5

Db 30 GTYVE 34

## RESULT 4

Q8XTM4 PRELIMINARY; PRT; 76 AA.  
 ID Q8XTM4  
 AC Q8XTM4;  
 DT 01-WAR-2002 (TRENBLrel. 20, Created)  
 DT 01-WAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Hypothetical protein RSP0082.  
 GN Name=R305556; OrderedLocusNames=RSP0082;  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).



```

OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler C., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguiet P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646076; CAD17233.1; -.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 76 AA; 8342 MW; 3673AE22EB1F9B39 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 76;
Best Local Similarity 80.0%; Pred. No. 8.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 70 GYAVE 74
|||

RESULT 5
DY21_ECOLI STANDARD; PRT; 78 AA.
AC P00383;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Dihydrofolate reductase type II (EC 1.5.1.3).
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84237581; PubMed=6735180;
RA Brieson N., Hohn T.;
RT "Nucleotide sequence of the dihydrofolate-reductase gene borne by the
RT plasmid R67 and conferring methotrexate resistance.";
RL Gene 28:271-275(1984).
RN [2]
RP SEQUENCE.
RX MEDLINE=80049683; PubMed=387759;
RA Stone D., Smith S.L.;
RT "The amino acid sequence of the trimethoprim-resistant dihydrofolate
RT reductase specified in Escherichia coli by R-plasmid R67.";
RL J. Biol. Chem. 254:10857-10861(1979).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 17-78.
RX MEDLINE=96069790; PubMed=7583655;
RA Narayana N., Matthews D.A., Howell E.E., Nguyen-Huu X.;
RT "A plasmid-encoded dihydrofolate reductase from trimethoprim-resistant
RT bacteria has a novel D2-symmetric active site.";
RL Nat. Struct. Biol. 2:1018-1025(1995).
CC -!- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADPH(+) = 7,8-
CC dihydrofolate + NADPH.
CC -!- PATHWAY: Essential step for de novo glycine and purine synthesis,
CC DNA precursor synthesis, and for the conversion of dUMP to dTMP.
CC -!- MISCELLANEOUS: Type II plasmid-specified enzyme is practically
CC insensitive to trimethoprim and methotrexate.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02118; AAA26083.1; -.
DR PIR; A91512; RDECD6.
DR PDB; 1VIE; X-ray; @=17-78.
DR PDB; 1VIF; X-ray; @=17-78.
DR InterPro; IPR009159; Dhfr_type_II.
DR InterPro; IPR008990; E_transp_acc.
DR Pfam; PF06442; DHFR; 1.
DR PIRSF; PIRSF000199; Dhfr_type_II; 1.
DR 3D-structure; Direct protein sequencing; Methotrexate resistance;
KW NADP; One-carbon metabolism; Oxidoreductase; Plasmid;
KW Trimethoprim resistance.
RN Trimmer 26 27
FT TURN 26 27
FT STRAND 29 32
FT STRAND 39 46
FT STRAND 54 59
FT TURN 63 64
FT STRAND 66 70
FT HELIX 71 73
FT STRAND 74 76
SQ SEQUENCE 78 AA; 8446 MW; 0BD80B9146529417 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 78;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 54 GYAVE 58
|||

RESULT 6
DY22_ECOLI STANDARD; PRT; 78 AA.
AC P00384;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Dihydrofolate reductase type II (EC 1.5.1.3).
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81174731; PubMed=6261228;
RA Zolg J.W., Haenggi U.J.;
RT "Characterization of a R plasmid-associated, trimethoprim-resistant
RT dihydrofolate reductase and determination of the nucleotide sequence
RT of the reductase gene.";
RL Nucleic Acids Res. 9:697-710(1981).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81269991; PubMed=7022127;
RA Swift G., McCarthy B.J., Heffron F.;
RT "DNA sequence of a plasmid-encoded dihydrofolate reductase.";
RL Mol. Gen. Genet. 181:441-447(1981).
CC -!- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADP(+) = 7,8-
CC dihydrofolate + NADPH.
CC -!- PATHWAY: Essential step for de novo glycine and purine synthesis,
CC DNA precursor synthesis, and for the conversion of dUMP to dTMP.
CC -!- MISCELLANEOUS: Type II plasmid-specified enzyme is practically
CC insensitive to trimethoprim and methotrexate.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00252; CAA23503.1; -.
DR EMBL; J01773; AAA72145.1; -.
DR EMBL; A00040; CAA00003.1; -.
DR PIR; A00398; RDECD8.
DR HSSP; P00383; IVIE.
DR InterPro; IPR009159; Dhfr_type_II.
DR InterPro; IPR008990; E_transp_acc.
DR Pfam; PF06442; DHFR; 1.
DR PIRSF; PIRSF000199; Dhfr type II; 1.
KW Methotrexate resistance; NADP; One-carbon metabolism; Oxidoreductase;
KW Plasmid; Trimethoprim resistance.
SQ SEQUENCE 78 AA; 8195 MW; EE9B686CBESBCEB93 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 78;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 54 GYAVE 58

RESULT 7
DY23_ECOLI STANDARD; PRT; 78 AA.
ID DY23_ECOLI
AC P05794;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Dihydrofolate reductase type II (EC 1.5.1.3).
OS Escherichia coli, and
OS Klebsiella aerogenes.
OG Plasmid IncP-beta R751.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 28451;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=86286598; PubMed=3526286;
RA Flensburg J., Steen R.;
RT "Nucleotide sequence analysis of the trimethoprim resistant
RT dihydrofolate reductase encoded by R plasmid R751.";
RL Nucleic Acids Res. 14:5933-5933 (1986).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=K.aerogenes;
RA Raderstrom P., Sundstroem L., Swedberg G., Flensburg J., Skoeld O.;
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=K.aerogenes;
RA Thomas C.M.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADP(+) = 7,8-
CC dihydrofolate + NADPH.
CC -1- PATHWAY: Essential step for de novo glycine and purine synthesis,
CC DNA precursor synthesis, and for the conversion of dUMP to dTMP.
CC -1- MISCELLANEOUS: Type II plasmid-specified enzyme is practically
CC insensitive to trimethoprim and methotrexate.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; X04128; CAA27740.1; -.
DR EMBL; X72585; CAA51181.1; -.
DR EMBL; U67194; AAC64446.1; -.
DR PIR; A23598; RDECD5.
DR PIR; S32183; S32183.
DR HSSP; P00383; IVIE.
DR InterPro; IPR009159; Dhfr_type_II.
DR InterPro; IPR008990; E_transp_acc.
DR Pfam; PF06442; DHFR; 1.
DR PIRSF; PIRSF000199; Dhfr type II; 1.
KW Methotrexate resistance; NADP; One-carbon metabolism; Oxidoreductase;
KW Plasmid; Trimethoprim resistance.
SQ SEQUENCE 78 AA; 8328 MW; 3D67C2EB67220697 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 78;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 54 GYAVE 58

RESULT 8
Q71RY0 PRELIMINARY; PRT; 78 AA.
ID Q71RY0
AC Q71RY0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Dihydrofolate reductase.
GN Name=dfrIIc;
OS Aeromonas salmonicida subsp. salmonicida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;
OC Aeromonadaceae; Aeromonas.
OX NCBI_TaxID=29491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=324-94 AS44;
RX MEDLINE=21594646; PubMed=11759088;
RA L'Abbe-Lund T.M., Sorum H.;
RT "Class 1 integrons mediate antibiotic resistance in the fish pathogen
RT Aeromonas salmonicida worldwide.";
RL Microb. Drug Resist. 7:263-272 (2001).
DR EMBL; AF327729; AAK53555.1; -.
DR InterPro; IPR009159; Dhfr_type_II.
DR InterPro; IPR008990; E_transp_acc.
DR Pfam; PF06442; DHFR; 2; 1.
DR PIRSF; PIRSF000199; Dhfr type II; 1.
DR SEQUENCE 78 AA; 8328 MW; 3D67C2EB67220697 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 78;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 54 GYAVE 58

RESULT 9
Q79NS3 PRELIMINARY; PRT; 78 AA.
ID Q79NS3
AC Q79NS3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Dihydrofolate reductase.
GN Name=dfrB2; bacterium.
OS uncultured pCEM5.
OG Bacteria; environmental samples.
OX NCBI_TaxID=77133;
```

```

RN SEQUENCE FROM N.A.
RP Tennstedt T., Szczepanowski R., Braun S., Puehler A., Schluter A.;
RT "Occurrence of integron-associated resistance gene cassettes located
RT on antibiotic resistance plasmids isolated from a wastewater treatment
RT plant.";
RL FEMS Microbiol. Ecol. 45:239-252(2003).
DR EMBL; AY139592; XAN41415.1; -.
DR InterPro; IPR009159; Dhfr_type_II.
DR InterPro; IPR008990; E_transp_acc.
DR Pfam; PF06442; DHFR_2; 1.
DR PIRSF; PIRSF000199; Dhfr_type_II; 1.
KW Plasmid.
SQ SEQUENCE 78 AA; 8195 MW; EE9B686CBESBCB33 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 78;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 54 GYAVE 58

RESULT 10
Q8VNN1 PRELIMINARY; PRT; 78 AA.
AC Q8VNN1;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Dihydrofolate reductase.
GN Name=dfr2d;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Grape M., Sundstroem L., Kronvall G.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ429132; CAD22098.1; -.
DR HSP; P00383; IVIE.
DR GO; GO:0004146; F:dihydrofolate reductase activity; IEA.
DR GO; GO:0042493; P:response to drug; IEA.
DR InterPro; IPR009159; Dhfr_type_II.
DR InterPro; IPR008990; E_transp_acc.
DR Pfam; PF06442; DHFR_2; 1.
DR PIRSF; PIRSF000199; Dhfr_type_II; 1.
SQ SEQUENCE 78 AA; 8410 MW; AE14E2131AB6BB4E CRC64;

Query Match 91.7%; Score 22; DB 2; Length 78;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 54 GYAVE 58

RESULT 11
Q6D6P6 PRELIMINARY; PRT; 78 AA.
AC Q6D6P6;
DT 01-OCT-2004 (TReMBLrel. 28, Created)
DT 01-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein.
ORFNames=ECA1636;
OS Erwinia carotovora subsp. atroseptica SCRI1043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=SCR11043;
RA Bell K.S., Sebathia M., Pritchard L., Holden M., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BX950851; CAG74540.1; -.
KW Hypothetical protein.
SQ SEQUENCE 78 AA; 8646 MW; 52B7EC35B09BDD9 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 78;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 35 GYSVE 39

RESULT 12
AAM89276 PRELIMINARY; PRT; 78 AA.
ID AAM89276;
AC AAM89276;
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE DfFB3/DHFR1c dihydrofolate reductase.
GN DFRB3.
OS Escherichia coli.
OG Plasmid pBG0100.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VA292;
RA Patridge S.R., Hall R.M.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY123252; AAM89276.1; -.
KW Plasmid.
SQ SEQUENCE 78 AA; 8328 MW; 3D67C2EB67220697 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 78;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 54 GYAVE 58

RESULT 13
AAA82255 PRELIMINARY; PRT; 78 AA.
ID AAA82255;
AC AAA82255;
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE Dihydrofolate reductase.
GN DFR2A.
OS Escherichia coli.
OG Plasmid pMO239.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=37A-2;
RA Hansson K., Sundstrom L.;

```

RT "The dfr2a trimethoprim resistance gene is born on an integron cassette.";  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U36276; AAA82255.1; -;  
 SQ SEQUENCE 78 AA; 8446 MW; 0BDB0B9146529417 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 78;  
 Best Local Similarity 80.0%; Pred. No. 8.9e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 DB 54 GYAVE 58

RESULT 14  
 ID AAK53555 PRELIMINARY; PRT; 78 AA.  
 AC AAK53555;  
 DT 02-MAR-2004 (TREMELrel. 27, Created)  
 DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)  
 DE Dihydrofolate reductase.  
 GN DFR1C.  
 OS Aeromonas salmonicida subsp. salmonicida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;  
 OC Aeromonas; Aeromonas salmonicida.  
 OX NCBI\_TaxID=29491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=324-94 AS44;  
 RA L'Abbe-Lund T.M., Sorum H.;  
 RT "Class 1 integrons mediate antibiotic resistance in the fish pathogen  
 RT Aeromonas salmonicida worldwide.";  
 RL Microb. Drug Resist. 7:263-272(2001).  
 DR EMBL; AF327729; AAK53555.1; -;  
 SQ SEQUENCE 78 AA; 8328 MW; 3D67C2EB67220697 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 78;  
 Best Local Similarity 80.0%; Pred. No. 8.9e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 DB 54 GYAVE 58

RESULT 15  
 ID AAK95982 PRELIMINARY; PRT; 78 AA.  
 AC AAK95982;  
 DT 02-MAR-2004 (TREMELrel. 27, Created)  
 DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)  
 DE Dihydrofolate reductase.  
 GN DFR2.  
 OS Escherichia coli.  
 OG Plasmid R386.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21956005; PubMed=11959558;  
 RA Partridge S.R., Brown H.J., Hall R.M.;  
 RT "Characterization and movement of the class 1 integron known as tn2521  
 RT and tn1405.";  
 RL Anticrib. Agents Chemother. 46:1288-1294(2002).  
 DR EMBL; U12441; AAK95982.1; -;  
 SQ Plasmid.  
 Query Match 91.7%; Score 22; DB 2; Length 78;

Best Local Similarity 80.0%; Pred. No. 8.9e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 DB 54 GYAVE 58

RESULT 16  
 ID AAN41415 PRELIMINARY; PRT; 78 AA.  
 AC AAN41415;  
 DT 02-MAR-2004 (TREMELrel. 27, Created)  
 DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)  
 DE Dihydrofolate reductase.  
 GN DFRB2.  
 OS Gamma-proteobacterium Hot 75m4.  
 OG Plasmid pCEm5.  
 OC Bacteria; environmental samples.  
 OX NCBI\_TaxID=77133;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Temstedt T., Szczepanowski R., Braun S., Puhler A., Schluter A.;  
 RT "Occurrence of integron-associated resistance gene cassettes located  
 RT on antibiotic resistance plasmids isolated from a wastewater treatment  
 RT plant.";  
 RL FEMS Microbiol. Ecol. 45:239-252(2003).  
 DR EMBL; AY139592; AAN41415.1; -;  
 KW Plasmid.  
 SQ SEQUENCE 78 AA; 8195 MW; EE9B686CBE5BCB93 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 78;  
 Best Local Similarity 80.0%; Pred. No. 8.9e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 DB 54 GYAVE 58

RESULT 17  
 ID O99013 PRELIMINARY; PRT; 81 AA.  
 AC O99013;  
 DT 01-MAY-1999 (TREMELrel. 10, Created)  
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
 DE ATP synthase CF0 C chain.  
 GN Name=atpH;  
 OS Proteobacteria wickerhamii.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
 OC Chlorellaceae; Prototheca.  
 OX NCBI\_TaxID=3111;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=263-11;  
 RA Knauf U., Hachtel W.;  
 RT "The genes encoding subunits of ATP synthase are conserved in the  
 RT reduced plastid genome of the heterotrophic alga Prototheca  
 RT wickerhamii.";  
 RL Mol. Genet. Genomics 267:492-497(2004).  
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic  
 CC core - and CF(0) the membrane proton channel (By similarity).  
 CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic  
 CC core - and CF(0) - the membrane proton channel. CF(1) has five  
 CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)  
 CC has three main subunits: a, b and c (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane (By  
 CC similarity).

```

CC -!- SIMILARITY: Belongs to the ATPase C chain family.
DR EMBL; AJ236874; CAB3451.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0016469; C:proton-transferring two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transferring ATP synthase activity; .; IEA.
DR GO; GO:0046961; F:hydrogen-transferring ATPase activity, rota. .; IEA.
DR GO; GO:0016820; F:hydrolyase activity, acting on acid anhydrid. .; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR002379; ATPase_Csub.
DR InterPro; IPR005953; ATP_synth_C.
DR InterPro; IPR000454; Eub_ATPase_Csub.
DR Pfam; PF00137; ATP-synt_C; 1.
DR PRINTS; PR00124; ATPASEC.
DR TIGRFAMs; TIGR01260; ATP_synt_c; 1.
DR PROSITE; PS00605; ATPASE_C; 1.
KW CF(0); Chloroplast; Direct protein sequencing; Formylation;
KW Lipid-binding; Transmembrane; Transport.
KW Lipid-binding; Transmembrane; Transport.
SQ SEQUENCE 81 AA; 8048 MW; 17324FDAF38278EE CRC64;

Query Match 91.7%; Score 22; DB 2; Length 81;
Best Local Similarity 80.0%; Pred. No. 9.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db ||||
33 GYAVE 37

RESULT 18
ID ATPH CHLRE STANDARD; PRT; 82 AA.
AC Q37304; Q9T2G4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein) (Subunit
DE III).
GN Name=atpH;
OS Chlamydomonas reinhardtii.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cw15;
RX MEDLINE=98031899; PubMed=9362486;
RA Rolland N., Dorne A.-J., Amoroso G., Sueltemeyer D.F., Joyard J.,
RA Rochaix J.-D.;
RT "Disruption of the plastid ycf10 open reading frame affects uptake of
RT inorganic carbon in the chloroplast of Chlamydomonas.";
RN EMBO J. 16:6713-6726 (1997).
RN [2]
RP SEQUENCE OF 1-32.
RC STRAIN=cw15;
RX MEDLINE=96128220; PubMed=8543042;
RA Fiedler H.K., Schmid R., Leu S., Shavit N., Strotmann H.;
RT "Isolation of CFOCF1 from Chlamydomonas reinhardtii cw15 and the N-
RT terminal amino acid sequences of the CFOCF1 subunits.";
RN FEBS Lett. 377:163-166 (1995).
RN [3]
RN COMPLETE PLASTID GENOME.
RX MEDLINE=22305394; PubMed=12417694;
RA Maul J.E., Lilly J.W., Cui L., defamphilis C.W., Miller W.,
RA Harris E.H., Stern D.B.;
RT "The Chlamydomonas reinhardtii plastid chromosome: islands of genes in
RT a sea of repeats.";
RL Plant Cell 14:2659-2679 (2002).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
H(+) (Out).

```

---

```

CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: a, b and c.
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- MISCELLANEOUS: Dicyclohexylcarbodiimide (DCCD) inhibits ATPase.
CC -!- SIMILARITY: Belongs to the ATPase C chain family.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X90559; CAA62149.1; -.
DR EMBL; BK000554; DAA00954.1; -.
DR PIR; S58349; S58349.
DR InterPro; IPR002379; ATPase_Csub.
DR InterPro; IPR005953; ATP_synth_C.
DR InterPro; IPR000454; Eub_ATPase_Csub.
DR Pfam; PF00137; ATP-synt_C; 1.
DR PRINTS; PR00124; ATPASEC.
DR TIGRFAMs; TIGR01260; ATP_synt_c; 1.
DR PROSITE; PS00605; ATPASE_C; 1.
KW CF(0); Chloroplast; Direct protein sequencing; Formylation;
KW Hydrogen ion transport; Lipid-binding; Transmembrane.
FT TRANSMEM 3 23 Potential.
FT TRANSMEM 57 77 Potential.
FT MOD_RES 1 1 N-formylmethionine.
SQ SEQUENCE 82 AA; 8092 MW; BFDC3BAF43E96A45 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 82;
Best Local Similarity 80.0%; Pred. No. 9.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db ||||
33 GYAVE 37

RESULT 19
ID ATPH CHLVU STANDARD; PRT; 82 AA.
AC P56297;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein) (Subunit
DE III).
GN Name=atpH;
OS Chlorella vulgaris (Green alga).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM C-27 / Tamiya;
RX MEDLINE=97303241; PubMed=9159184;
RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Inamura A., Yoshinaga K., Sugitara M.;
RT "Complete nucleotide sequence of the chloroplast genome from the green
RT alga Chlorella vulgaris: the existence of genes possibly involved in
RT chloroplast division.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972 (1997).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)

```

```
CC      has three main subunits: a, b and c.
CC      -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC      -!- MISCELLANEOUS: Dicyclohexylcarbodiimide (DCCD) inhibits ATPase.
CC      -!- SIMILARITY: Belongs to the ATPase C chain family.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; AB001684; BAA57858.1; -.
CC      PIR; T07211; T07211.
CC      InterPro; IPR002379; ATPase_Caeb.
CC      InterPro; IPR005953; ATP synth C.
CC      InterPro; IPR000454; Eub_ATPase_Caeb.
CC      Pfam; PF00137; ATP-synt_C; 1.
CC      PRINTS; PR00124; ATPASEC.
CC      TIGRFAMs; TIGR01260; ATP_synt_c; 1.
CC      PROSITE; PS00605; ATPASE C; 1.
CC      CF(0); Chloroplast; Hydrogen ion transport; Lipid-binding;
KW      Transmembrane.
FT      TRANSMEM 3 23 Potential.
FT      TRANSMEM 57 77 Potential.
FT      TRANSMEM 82 AA; 8060 MW; 3569FFDFE1P9632E CRC64;
SQ      SEQUENCE 82 AA; 8060 MW; 3569FFDFE1P9632E CRC64;

Query Match 91.7%; Score 22; DB 1; Length 82;
Best Local Similarity 80.0%; Pred. No. 9.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      33 GYAVE 37

RESULT 20
Q8FTB1 PRELIMINARY; PRT; 83 AA.
AC Q8FTB1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein MW2806.
GN OrderedLocusNames=WM2806;
OS Methanosarcina mazei (Methanosarcina frisia)
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / GoI / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wieser A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RA "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013530; AM32502.1; -.
DR InterPro; IPR003850; Pufs.
DR Pfam; PF02700; Pufs; 1.
DR ProDom; PD010362; Pufs; 1.
DR TIGRFAMs; TIGR00302; UPF0062; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 83 AA; 9283 MW; 8FE3CE5632B224 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 83;
Best Local Similarity 80.0%; Pred. No. 9.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CC      has three main subunits: a, b and c.
CC      -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC      -!- MISCELLANEOUS: Dicyclohexylcarbodiimide (DCCD) inhibits ATPase.
CC      -!- SIMILARITY: Belongs to the ATPase C chain family.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; AB001684; BAA57858.1; -.
CC      PIR; T07211; T07211.
CC      InterPro; IPR002379; ATPase_Caeb.
CC      InterPro; IPR005953; ATP synth C.
CC      InterPro; IPR000454; Eub_ATPase_Caeb.
CC      Pfam; PF00137; ATP-synt_C; 1.
CC      PRINTS; PR00124; ATPASEC.
CC      TIGRFAMs; TIGR01260; ATP_synt_c; 1.
CC      PROSITE; PS00605; ATPASE C; 1.
CC      CF(0); Chloroplast; Hydrogen ion transport; Lipid-binding;
KW      Transmembrane.
FT      TRANSMEM 3 23 Potential.
FT      TRANSMEM 57 77 Potential.
FT      TRANSMEM 82 AA; 8060 MW; 3569FFDFE1P9632E CRC64;
SQ      SEQUENCE 82 AA; 8060 MW; 3569FFDFE1P9632E CRC64;

Query Match 91.7%; Score 22; DB 1; Length 82;
Best Local Similarity 80.0%; Pred. No. 9.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      33 GYAVE 37

RESULT 20
Q8FTB1 PRELIMINARY; PRT; 83 AA.
AC Q8FTB1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein MW2806.
GN OrderedLocusNames=WM2806;
OS Methanosarcina mazei (Methanosarcina frisia)
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / GoI / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wieser A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RA "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013530; AM32502.1; -.
DR InterPro; IPR003850; Pufs.
DR Pfam; PF02700; Pufs; 1.
DR ProDom; PD010362; Pufs; 1.
DR TIGRFAMs; TIGR00302; UPF0062; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 83 AA; 9283 MW; 8FE3CE5632B224 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 83;
Best Local Similarity 80.0%; Pred. No. 9.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      33 GYAVE 37

RESULT 20
Q8FTB1 PRELIMINARY; PRT; 83 AA.
AC Q8FTB1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein MW2806.
GN OrderedLocusNames=WM2806;
OS Methanosarcina mazei (Methanosarcina frisia)
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / GoI / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wieser A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RA "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013530; AM32502.1; -.
DR InterPro; IPR003850; Pufs.
DR Pfam; PF02700; Pufs; 1.
DR ProDom; PD010362; Pufs; 1.
DR TIGRFAMs; TIGR00302; UPF0062; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 83 AA; 9283 MW; 8FE3CE5632B224 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 84;
Best Local Similarity 80.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      32 GYSVE 36

RESULT 22
YE86 CLOAB STANDARD; PRT; 86 AA.
ID YE86 CLOAB
AC Q97J00;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical UPF0180 protein CAC1486.
GN OrderedLocusNames=CAC1486;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RX DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Tatusov R.L., Sabathe F., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -!- SIMILARITY: Belongs to the UPF0180 family.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
```

```
QY      1 GYXVE 5
DB      31 GYAVE 35

RESULT 21
Q8ZZJ8 PRELIMINARY; PRT; 84 AA.
ID Q8ZZJ8;
AC Q8ZZJ8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PAE0224.
GN OrderedLocusNames=PAE0224;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RA "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009756; AAL62641.1; -.
DR InterPro; IPR003850; Pufs.
DR ProDom; PD010362; Pufs; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 84 AA; 9121 MW; A81188CBD3052C72 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 84;
Best Local Similarity 80.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      32 GYSVE 36

RESULT 22
YE86 CLOAB STANDARD; PRT; 86 AA.
ID YE86 CLOAB
AC Q97J00;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical UPF0180 protein CAC1486.
GN OrderedLocusNames=CAC1486;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RX DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Tatusov R.L., Sabathe F., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -!- SIMILARITY: Belongs to the UPF0180 family.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF007659; AAK79454.1; --  
 DR PIR; C97083; C97083.  
 DR HAWAP; MF\_00506; --; 1.  
 DR InterPro; IPR005370; UPF0180.  
 DR Pfam; PF03698; UPF0180; 1.  
 DR ProDom; PD059670; UPF0180; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 86 AA; 9299 MW; A5DFED6283B6A4E1 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 86;  
 Best Local Similarity 80.0%; Pred. No. 9.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
 ||||  
 Db 21 GYSVE 25

## RESULT 23

Q8KK08 PRELIMINARY; PRT; 86 AA.  
 AC Q8KK08;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein orf178.  
 GN Name=orf178;  
 OS Proteus vulgaris.  
 OG Plasmid Rts1.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Proteus.  
 OX NCBI\_TaxID=585;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX PubMed=12029035;  
 RA Murata T., Ohnishi M., Ara T., Kaneko J., Han C.-G., Li Y.F.,  
 RA Takashima K., Nojima H., Nakayama K., Kaji A., Kamio Y., Miki T.,  
 RA Mori H., Ohtsubo E., Terawaki Y., Hayashi T.;  
 RT "Complete nucleotide sequence of plasmid Rts1: implications for  
 RT evolution of large plasmid genomes.";  
 RL J. Bacteriol. 184:3194-3202(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX MEDLINE=96184644; PubMed=8645296;  
 RA Tian Q.B., Ohnishi M., Tabuchi A., Terawaki Y.;  
 RT "A new plasmid-encoded proteic killer gene system: cloning,  
 RT sequencing, and analyzing hig locus of plasmid Rts1.";  
 RL Biochem. Biophys. Res. Commun. 220:280-284(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX MEDLINE=9435883; PubMed=8078071;  
 RA Janosi L., Yonemitsu H., Hong H., Kaji A.;  
 RT "Molecular cloning and expression of a novel hydroxymethylcytosine-  
 RT specific restriction enzyme (PvuRts1) modulated by glucosylation of  
 RT DNA.";  
 RL J. Mol. Biol. 242:45-61(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX MEDLINE=91193219; PubMed=2013575;  
 RA Mochida S., Tsuchiya H., Mori K., Kaji A.;  
 RT "Three short fragments of Rts1 DNA are responsible for the  
 RT temperature-sensitive growth phenotype (tsg) of host bacteria.";  
 RL J. Bacteriol. 173:2600-2607(1991).  
 RN [5]

RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX PubMed=2840681;  
 RA Nozue H., Tsuchiya K., Kamio Y.;  
 RT "Nucleotide sequence and copy control function of the extension of the  
 RT incl region (incl-b) of Rts 1.";  
 RL Plasmid 19:46-56(1988).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX MEDLINE=88139175; PubMed=3277947;  
 RA Tanaka M., Okawa N., Mori K., Suyama Y., Kaji A.;  
 RT "Nucleotide sequence of an Rts1 fragment causing temperature-dependent  
 RT instability.";  
 RL J. Bacteriol. 170:1175-1182(1988).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX MEDLINE=85234397; PubMed=2989253;  
 RA Mollet B., Clerget M., Meyer J., Iida S.;  
 RT "Organization of the Tn6-related kanamycin resistance transposon  
 RT Tn2680 carrying two copies of IS26 and an IS903 variant, IS903. B.";  
 RL J. Bacteriol. 163:55-60(1985).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX MEDLINE=84185439; PubMed=6325393;  
 RA Kamio Y., Tabuchi A., Itoh Y., Katagiri H., Terawaki Y.;  
 RT "Complete nucleotide sequence of mini-Rts1 and its copy mutant.";  
 RL J. Bacteriol. 158:307-312(1984).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX MEDLINE=83290717; PubMed=6309744;  
 RA Kamio Y., Terawaki Y.;  
 RT "Nucleotide sequence of an incompatibility region of mini-Rts1 that  
 RT contains five direct repeats.";  
 RL J. Bacteriol. 155:1185-1191(1983).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UR-75;  
 RX MEDLINE=68393387; PubMed=4876466;  
 RA Terawaki Y., Kakizawa Y., Takayasu H., Yoshikawa M.;  
 RT "Temperature sensitivity of cell growth in Escherichia coli associated  
 RT with the temperature sensitive R(KM) factor.";  
 RL Nature 219:284-285(1968).  
 DR EMBL; AP004237; BAB93740.1; --  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 86 AA; 9392 MW; 45B9D72C69EC2B2D CRC64;  
 Query Match 91.7%; Score 22; DB 2; Length 86;  
 Best Local Similarity 80.0%; Pred. No. 9.8e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 GYXVE 5  
 ||||  
 Db 19 GYAVE 23

## RESULT 24

Q6D329 PRELIMINARY; PRT; 87 AA.  
 ID Q6D329;  
 AC Q6D329;  
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=ECA2915;  
 OS Erwinia carotovora subsp. atroseptica SCRI1043.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Pectobacterium.  
 OX NCBI\_TaxID=218491;  
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-SCRI1043;
RA Bell K.S., Sebathia M., Pritchard L., Holden M., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Atkin R., Bacon N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.;
RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BX950851; CAG75815.1; -.
RW Hypothetical protein.
SQ SEQUENCE 87 AA; 9776 MW; 64E1C62656890DD6E CRC64;

Query Match 91.7%; Score 22; DB 2; Length 87;
Best Local Similarity 80.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 44 GYSVE 48

RESULT 25
Q8TPE9 PRELIMINARY; PRT; 88 AA.
AC Q8TPE9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein MA1964.
GN OrderedLocusNames=MA1964;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Kincor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R.,
RA Linton L., McSwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.B., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010880; AAM05367.1; -.
DR InterPro; IPR003850; Pufs.
DR Pfam; PF02700; PurC; 1.
DR ProDom; PD010362; Pufs; 1.
DR TIGRFAMs; TIGR00302; UPF0062; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 88 AA; 9850 MW; 2AA7DAC0F43FD379 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 88;
Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 36 GYAVE 40

RESULT 26
Q48265 PRELIMINARY; PRT; 91 AA.
AC Q48265;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Outer membrane P6 protein homologue (Fragment).
OS Haemophilus parainfluenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=729;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueyama T., Kuroda Y., Shirabe K., Takeshita M., Mogi G.;
RA Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RL EMBL; D28887; BAA06035.1; -.
CC -1- SIMILIARITY: Belongs to the ompA family.
DR GO; GO:0016021; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR GO; GO:0009279; C:integral to membrane; IEA.
DR InterPro; IPR006664; Bac_OmpA.
DR InterPro; IPR006665; OmpA/MotB.
DR Pfam; PF00691; OmpA; 1.
DR PRINTS; PR01021; OMPADOMAIN.
DR ProDom; PD000930; OmpA/MotB; 1.
FT NON_TER 1 1
FT NON_TER 91 91
SQ SEQUENCE 91 AA; 9896 MW; C6A6FCACB5CE434 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 91;
Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 4 GYSVE 8

RESULT 27
Q982Z2 PRELIMINARY; PRT; 94 AA.
AC Q982Z2;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transcriptional regulator.
GN OrderedLocusNames=mlr8432;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082936; PubMed=11214974;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti (supplement).";
RL DNA Res. 7:381-406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AP003014; BAB54314.1; -.

```



```

DR GO: 0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
DR PROSITE; PS00943; HTH_CROCI; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 94 AA; 9895 MW; 02CD173EB020CF25 CRC64;

Query Match          91.7%; Score 22; DB 2; Length 94;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 33 GYSVE 37

RESULT 28
YAYB_CALSA
ID YAYB_CALSA STANDARD; PRT; 97 AA.
AC P23554;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 10.7 kDa protein in xynB 3'region (ORF 3).
OS Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
OC Caldicellulosiruptor.
OX NCBI_TaxID=44001;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90253140; PubMed=2111111;
RA Luethi E., Love D.R., McAnulty J., Wallace C., Caughey P.A., Saul D.,
RA Bergquist P.L.;
RT "Cloning, sequence analysis, and expression of genes encoding xylan-
RT degrading enzymes from the thermophile 'Caldocellum
RT saccharolyticum'.";
RL Appl. Environ. Microbiol. 56:1017-1024 (1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34459; AAA23061.1; -.
DR PIR; C37202; C37202.
KW Hypothetical protein.
SQ SEQUENCE 97 AA; 10731 MW; CEF79D73BD2884F3 CRC64;

Query Match          91.7%; Score 22; DB 1; Length 97;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 69 GYSVE 73

RESULT 29
Q8GL63
ID Q8GL63 PRELIMINARY; PRT; 97 AA.
AC Q8GL63;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Dihydrofolate reductase.
GN Name=dfriI;
OS uncultured bacterium.
OG Plasmid pSp39.

Bacteria; environmental samples.
OC NCBI_TaxID=77133;
RN [1]
RP SEQUENCE FROM N.A.
RA Tennstedt T., Szczepanowski R., Braun S., Puehler A., Schluster A.;
RT "Occurrence of integron-associated resistance gene cassettes located
RT on antibiotic resistance plasmids isolated from a wastewater treatment
RT plant.";
RL FEMS Microbiol. Ecol. 45:239-252 (2003).
DR EMBL; AY139601; AAN41433.1; -.
DR HSSP; P00383; 1VIE.
DR GO: 0004146; P:dihydrofolate reductase activity; IEA.
DR GO: 0042493; P:response to drug; IEA.
DR InterPro; IPR009159; Dhfr_type_II.
DR InterPro; IPR008990; E_transp_acc.
DR Pfam; PF06442; DHFR_2; 1.
DR PIRSF; PIRSF000199; Dhfr_type_II; 1.
KW Plasmid.
SQ SEQUENCE 97 AA; 10666 MW; F3080D432F9BF1E8 CRC64;

Query Match          91.7%; Score 22; DB 2; Length 97;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 73 GYAVE 77

RESULT 30
Q8IKH6
ID Q8IKH6 PRELIMINARY; PRT; 99 AA.
AC Q8IKH6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF14_0629;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511 (2002).
DR EMBL; AE014826; AAN37242.1; -.
KW Hypothetical protein.
SQ SEQUENCE 99 AA; 11305 MW; 7735A8ED088500C8 CRC64;

Query Match          91.7%; Score 22; DB 2; Length 99;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 22 GYTVE 26

RESULT 31
Q7MQT5
ID Q7MQT5 PRELIMINARY; PRT; 101 AA.
AC Q7MQT5;

```

DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein.  
DE OrderedLocusNames=WS2027;  
GN Wolinella succinogenes;  
OS Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Helicobacteriaceae; Wolinella.  
OX NCBI\_TaxID=844;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSMZ 1740;  
RX MEDLINE=22882897; PubMed=14500908;  
RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O., B.,  
RA Nandakumar R., Gross R., Rosinus A., Keller H., Dagtap P., Linke B.,  
RA Meyer F., Lederer H., Schuster S.C.;  
RT "Complete genome sequence and analysis of Wolinella succinogenes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).  
DR EMBL; BX571662; CAE11028.1;  
SQ SEQUENCE 101 AA; 10406 MW; D83E200D3A55F76F CRC64;  
Query Match 91.7%; Score 22; DB 2; Length 101;  
Best Local Similarity 80.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
Db 50 GYTVE 54  
RESULT 32  
YGSN\_YEAST STANDARD; PRT; 108 AA.  
ID YGSN\_YEAST  
AC P40326;  
DT 01-PEB-1995 (Rel. 31, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Hypothetical 12.3 kDa protein in P0L2-YTA7 intergenic region.  
GN OrderedLocusNames=YGR269W;  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=95274317; PubMed=7754704;  
RA Schnall R., Mannhaupt G., Stucka R., Tauer R., Ehnlé S.,  
RA Schwarzlöse C., Vetter I., Feldmann H.;  
RT "Identification of a set of yeast genes coding for a novel family of  
RT putative ATPases with high similarity to constituents of the 26S  
RT protease complex.";  
RL Yeast 10:1141-1155(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / FY1679;  
RA Agostoni Carbone M.L., Lucchini G., Melchiorretto P., Nardese V.,  
RA Varoni M., Panzeri L.;  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X81072; CAA56962.1; --  
DR EMBL; Y07893; CAA69200.1; --  
DR EMBL; Z73053; CAA97299.1; --  
DR PIR; S64602; S64602.

DR GermOnline; 141581; --  
DR SGD; S0003501; YGR269W.  
KW Hypothetical protein.  
FT CONFLICT 53 V -> A (in Ref. 1).  
SQ SEQUENCE 108 AA; 12279 MW; 660A0C63B87FC941 CRC64;  
Query Match 91.7%; Score 22; DB 1; Length 108;  
Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
Db 44 GYTVE 48  
RESULT 33  
QYX3H2 PRELIMINARY; PRT; 108 AA.  
ID QYX3H2  
AC QYX3H2;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=RB49ORF270C;  
OS Enterobacteria phage RB49.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC T4-like viruses.  
OX NCBI\_TaxID=50948;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97250366; PubMed=9096222;  
RA Monod C., Repolla F., Kutateladze M., Tetart F., Kriisch H.M.;  
RT "The genome of the pseudo T-even bacteriophages, a diverse group that  
RT resembles T4.";  
RL J. Mol. Biol. 267:237-249(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21062343; PubMed=11092834;  
RA Ang D., Keppel F., Klein G., Richardson A., Georgopoulos C.;  
RT "Genetic analysis of bacteriophage-encoded cochaparons.";  
RL Annu. Rev. Genet. 34:439-456(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21972795; PubMed=11976309;  
RA Desplats C., Dez C., Tetart F., Eleaume H., Kriisch H.M.;  
RT "Snapshot of the genome of the pseudo-T-even bacteriophage RB49.";  
RL J. Bacteriol. 184:2789-2804(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Ang D., Richardson A., Mayer M.P., Keppel F., Kriisch H.,  
RA Georgopoulos C.;  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Desplats C., Dez C., Tetart F., Eleaume H., Kriisch H.M.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Desplats C., Dez C., Tetart F., Eleaume H., Kriisch H.M.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RA Tetart F., Desplats C., Kutateladze M., Monod C., Ackermann H.-W.,  
RA Kriisch H.M.;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP SEQUENCE FROM N.A.  
RA Kriisch H.M.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [9]  
RP SEQUENCE FROM N.A.  
RA Zhao L., Tetart F., Kriisch H.M., Arisaka F.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

```

[10]
RN SEQUENCE FROM N.A.
RA Thiemer C.A., Desplats C., Dez C., Tetart F., Eleaume H., Krusch H.M.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[11]
RN SEQUENCE FROM N.A.
RA Krusch H.M.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
[12]
RN SEQUENCE FROM N.A.
RA Desplats C., Krusch H.M.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
[13]
RN SEQUENCE FROM N.A.
RA Letarov A.V., Krusch H.M., Tetart P.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
[14]
RN SEQUENCE FROM N.A.
RA Bertrand C., Petrov V., Nolan J., Letarov A., Desplats C., Chin D.,
Karam J.D., Krusch H.M.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY343333; AAQ15474.1; -.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 12244 MW; 8CC91182224B6616 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 108;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 103 GYTVE 107

RESULT 34
Q6FFPA4 PRELIMINARY; PRT; 108 AA.
AC Q6FFPA4;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative DNA binding protein.
GN OrderedLocusNames=ACIAD0291;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1] SEQUENCE FROM N.A.
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
FT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 0:0-0(2004).
DR EMBL; CR543861; CAG67253.1; -.
DR InterPro; IPR001801; Histone_HNS.
DR Pfam; PF00816; Histone_HNS; 1.
DR ProDom; PD007337; Histone_HNS; 1.
DR SMART; SM00528; HNS; 1.
KW Complete proteome.
SQ SEQUENCE 108 AA; 12531 MW; 80556DC26D1CA699 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 108;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 47 GYTVE 51

RESULT 35
[10]
RN SEQUENCE FROM N.A.
RA Marischky G., Rolfs A., Richardson A., Kane M., Baqui M., Taycher E.,
Hu Y., Vannberg F., Weger J., Kramer J., Moreira D., Kelley F.,
Zuo D., Raphael J., Hogle C., Jepson D., Williams J., Camargo A.,
Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E.,
LaBaer J.;
RT "Creation of the YFLEX clone resource: cloning of Saccharomyces
cerevisiae ORFs in the Gateway recombinational cloning system.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY558383; AAS56709.1; -.
SQ SEQUENCE 108 AA; 12279 MW; 660A0C63B87FC941 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 108;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 44 GYTVE 48

RESULT 36
O58646 PRELIMINARY; PRT; 109 AA.
AC O58646;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein PH0887.
GN OrderedLocusNames=PH0887;
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000004; BAA29981.1; -.
DR FIR; G71077; G71077.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 109 AA; 12419 MW; A22AC8142429D5E5 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 109;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 21 GYTVE 25
```

```
RESULT 37
Q7PFV9 PRELIMINARY; PRT; 112 AA.
AC Q7PFV9;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ENSANGP0000023342 (Fragment).
GN Name=ENSANGG0000021327;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008846; EAA45166.1; -.
FT NON TER 1
SQ SEQUENCE 112 AA; 12027 MW; 0E814A25023252C1 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 112;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 9 GYTVE 13

RESULT 38
Q6N3A0 PRELIMINARY; PRT; 114 AA.
AC Q6N3A0;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RP3794;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572605; CAE29235.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 114 AA; 11821 MW; 9627AD2A3564ED63 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 114;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 77 GYAVE 81

RESULT 39
CAE29235
Q7PFV9 PRELIMINARY; PRT; 112 AA.
AC Q7PFV9;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ENSANGP0000023342 (Fragment).
GN Name=ENSANGG0000021327;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008846; EAA45166.1; -.
FT NON TER 1
SQ SEQUENCE 112 AA; 12027 MW; 0E814A25023252C1 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 112;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 9 GYTVE 13

RESULT 40
Q979N5 PRELIMINARY; PRT; 115 AA.
AC Q979N5;
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE TVG1156444 protein.
GN Name=TVG1156444; OrderedLocusNames=TV1125;
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GSSI / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima T., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Kuno T., Yamamoto Y., Atamaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of thermoplasma volcanium.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
DR EMBL; AP000995; BAB60267.1; -.
KW Complete proteome.
SQ SEQUENCE 115 AA; 13233 MW; 6CCC4938602AADD5 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 115;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 108 GYSVE 112

RESULT 41
Q84XP9
```

```

ID Q84XP9 PRELIMINARY; PRT; 115 AA.
AC Q84XP9;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Putative purple acid phosphatase (Fragment).
OS Brassica rapa subsp. pekinensis (Chinese cabbage) (Celery cabbage).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=51351;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee D.-S., Song H., Hur Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY185355; AAO32057.1; -.
DR HSSP; P80366; 4KBP.
FT NON TER 1
SQ SEQUENCE 115 AA; 13022 MW; E075DEDBE27A29D5 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 115;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db ||||
89 GYAVE 93

RESULT 42
Q7CTV2 PRELIMINARY; PRT; 119 AA.
ID Q7CTV2;
AC Q7CTV2;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE AGR_L_1677p.
OS OrderedLocusNames=AGR_L_1677;
GN Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21608551; PubMed=11743194;
RX Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328 (2001).
DR EMBL; AE008282; AAK89412.1; -.
DR InterPro; IPR001387; HTH 3.
DR InterPro; IPR010982; Lambda_like_DNA.
DR SMART; SM00530; HTH_XRE; 1.
SQ SEQUENCE 119 AA; 12789 MW; BD7ED4DB177E0D6E CRC64;

Query Match 91.7%; Score 22; DB 2; Length 119;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db ||||
73 GYTVE 77

RESULT 43
Q28647 PRELIMINARY; PRT; 120 AA.
ID Q28647

```

```

AC Q28647;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein AF1626.
GN OrderedLocusNames=AF1626;
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=22234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M.L., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Glodek A.,
RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,
RA Zhou L., Overbeek R., Occayne J.D., Weidman J.F., McDonald L.A.,
RA Utterback T.R., Cotton M.D., Spriggs T., Artiach P., Kaine B.P.,
RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,
RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,
RA Woese C.R., Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370 (1997).
DR EMBL; AE000990; AAB89627.1; -.
DR PIR; A69453; A69453.
DR TIGR; AF1626; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 120 AA; 13040 MW; F35A1EBBC0E8F870 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 120;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db ||||
19 GYAVE 23

RESULT 44
Q9AAK1 PRELIMINARY; PRT; 122 AA.
ID Q9AAK1;
AC Q9AAK1;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Chemotaxis protein CheXIV.
GN OrderedLocusNames=CC0596;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolony J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
DR EMBL; AE005734; AAK22582.1; -.
DR PIR; B87323; 1JBE6.
DR HSSP; P06143; 1JBE6.

```

TIGR; CC0596; --  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0000156; F:two-component response regulator activity; IEA.  
 DR GO; GO:0007600; P:sensory perception; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.  
 DR InterPro; IPR011006; Chef\_like.  
 DR InterPro; IPR001789; Response\_reg.  
 DR Pfam; PF00072; Response\_reg; 1.  
 DR ProDom; PD000039; Response\_reg; 1.  
 DR SMART; SM00448; REC; 1.  
 DR PROSITE; PS0110; RESPONSE\_REGULATORY; 1.  
 DR Complete proteome; Phosphorylation; Sensory transduction.  
 SQ SEQUENCE 122 AA; 13391 MW; A48FAD9D9886D7DD CRC64;

Query Match 91.7%; Score 22; DB 2; Length 122;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+03; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 25 GYAVE 29

RESULT 45  
 Q9C7K1 PRELIMINARY; PRT; 123 AA.  
 AC Q9C7K1;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein F14G9.13.  
 GN Name=F14G9.13;  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,  
 RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,  
 RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC069159; AAG50904.1; --  
 DR FIR; C96604; C96604.  
 KW Hypothetical protein.  
 SQ SEQUENCE 123 AA; 13953 MW; 79A58071DE501A1C CRC64;

Query Match 91.7%; Score 22; DB 2; Length 123;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+03; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 38 GYSVE 42

RESULT 46  
 Q6NME7 PRELIMINARY; PRT; 127 AA.  
 ID Q6NME7  
 AC Q6NME7  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE At1g56260.  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Carninci P., Hayashizaki Y.,

RA Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,  
 RA Seki M., Shinozaki K., Ecker J.R.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BT011713; AAS49076.1; --  
 SQ SEQUENCE 127 AA; 14378 MW; F67E791AAA30A2B0 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 127;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+03; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 38 GYSVE 42

RESULT 47  
 AAS49076 PRELIMINARY; PRT; 127 AA.  
 ID AAS49076  
 AC AAS49076;  
 DT 10-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 10-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 10-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE At1g56260.  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Carninci P., Hayashizaki Y.,  
 RA Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,  
 RA Seki M., Shinozaki K., Ecker J.R.;  
 RA "Arabidopsis ORF clones."  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BT011713; AAS49076.1; --  
 SQ SEQUENCE 127 AA; 14378 MW; F67E791AAA30A2B0 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 127;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+03; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 38 GYSVE 42

RESULT 48  
 Q14556 PRELIMINARY; PRT; 129 AA.  
 ID Q14556  
 AC Q14556;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE DNA/endogenous human papillomavirus type 16 (HPV) DNA, right flank and  
 DE viral host junction. (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cervical carcinoma;  
 RX MEDLINE=87141358; PubMed=3029430;  
 RA Baker C.C., Phelps W.C., Lindgren V., Braun M.J., Gonda M.A.,  
 RA Howley P.M.;  
 RT "Structural and transcriptional analysis of human papillomavirus type  
 RT 16 sequences in cervical carcinoma cell lines."  
 RL J. Virol. 61:962-971(1987).  
 DR EMBL; M15780; AAA8082.1; --  
 DR HSSP; P03120; 1DIO.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0006275; P:regulation of DNA replication; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001866; E2 N.

DR Pfam; PF00508; PFV E2 N; 1.

DR ProDom; PD000678; E2 N; 1.

FT NON TER 1

SQ SEQUENCE 129 AA; 15049 MW; 68BF776CCB02B4A1 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 129;

Best Local Similarity 80.0%; Pred. NO. 1.5e+03;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 114 GYTV 118

RESULT 49

RS6E HALN1

ID RS6E HALN1 STANDARD; PRT; 131 AA.

AC Q9HJ5;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE 30S ribosomal protein S6e.

GN Name=rs6e; OrderedLocustNames=VNG2514G;

OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).

OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;

OC Halobacteriaceae; Halobacterium.

OX NCBI\_TaxID=64091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20504483; PubMed=1101950; DOI=10.1073/pnas.190337797;

RA NG W.V., Kennedy S.P., Mahairs G.G., Berquist B., Pan M.,

RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,

RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,

RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,

RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,

RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,

RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,

RA Ebbardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,

RT "Genome sequence of Halobacterium species NRC-1";

RL Proc. Natl. Acad. Sci. U.S.A. 97:12116-12181(2000).

CC -!- SIMILARITY: Belongs to the S6E family of ribosomal proteins.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AE005129; AAG20576.1; --

DR F01; D84401; D84401.

DR HAWAP; MF\_00512; --; 1.

DR InterPro; IPR001377; Ribosomal S6E.

DR Pfam; PF01092; Ribosomal S6e; 1.

DR PROSITE; PS00578; RIBOSOMAL S6E; 1.

KW Complete proteome; Ribosomal protein.

SQ SEQUENCE 131 AA; 13743 MW; 52FF77B5E6E2004B CRC64;

Query Match

Best Local Similarity 91.7%; Score 22; DB 1; Length 131;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 47 GYTV 51

RESULT 50

Q8C9X7

ID Q8C9X7 PRELIMINARY; PRT; 133 AA.

AC

DT

DT

DT

DE

DE

OS

OC

OC

OX

RN

RP

RC

RX

RA

RT

RL

RP

RC

RA

RA

RT

RL

RP

RC

RA

RA

RT

RL

RP

RC

RA

RA

RT

RL

RP

RC

RA

RA

RT

RL

RP

RC

RA

RA

RT

RL

RP

RC

RA

RA

RT

RL

RP

RC

RA

RA

RT

RL

RP

RC

RA

RA

RT

RL

RP

RC

RA

RA

RT

RL

RP

RC

RA

RA

Q8C9X7;

01-MAR-2003 (TrEMBLrel. 23, Created)

01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A430080K08 product:hypothetical protein, full insert sequence.

Mus musculus (Mouse).

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

NCBI\_TaxID=10090;

[1]

SEQUENCE FROM N.A.

STRAIN=C57BL/6J; TISSUE=Thymus;

MEDLINE=99279253; PubMed=10349636;

Carninci P., Hayashizaki Y.;

"High-efficiency full-length cDNA cloning.";

Meth. Enzymol. 303:19-44(1999).

[2]

SEQUENCE FROM N.A.

STRAIN=C57BL/6J; TISSUE=Thymus;

MEDLINE=21085660; PubMed=11217851;

RIKEN FANTOM Consortium;

"Functional annotation of a full-length mouse cDNA collection.";

[3]

SEQUENCE FROM N.A.

STRAIN=C57BL/6J; TISSUE=Thymus;

The FANTOM Consortium;

the RIKEN Genome Exploration Research Group Phase I & II Team;

"Analysis of the mouse transcriptome based on functional annotation of

60,770 full-length cDNAs.";

Nature 420:563-573(2002).

[4]

SEQUENCE FROM N.A.

STRAIN=C57BL/6J; TISSUE=Thymus;

MEDLINE=20499374; PubMed=11042159;

Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,

Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

"Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes.";

Genome Res. 10:1617-1630(2000).

[5]

SEQUENCE FROM N.A.

STRAIN=C57BL/6J; TISSUE=Thymus;

MEDLINE=20530913; PubMed=11076861;

Shibata K., Itoh M., Azawa K., Nagaoka S., Sasaki N., Carninci P.,

Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,

Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,

Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

"RIKEN integrated sequence analysis (RISA) system-384-format

sequencing pipeline with 384 multicapillary sequencer.";

Genome Res. 10:1757-1771(2000).

[6]

SEQUENCE FROM N.A.

STRAIN=C57BL/6J; TISSUE=Thymus;

Adachi J., Azawa K., Akimura T., Arakawa T., Bono H., Carninci P.,

Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,

Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,

Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,

Kurihara C., Matsuura T., Miyazaki A., Murata M., Nakamura M.,

Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,

Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,

Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,

Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,

Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

EMBL; AK040251; BAC30551.1; --

Hypothetical protein.

KW

Best Local Similarity 80.0%; Pred.No.1.5e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 31 GYSVE 35

RESULT 53  
Q9CYV3  
ID Q9CYV3 PRELIMINARY; PRT; 134 AA.  
AC Q9CV3;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Mus musculus 10, 11 days embryo whole body cDNA. RIKEN full-length  
DE enriched library. Clone:2B10440D03 product:10, 11 DAYS EMBRYO CDNA,  
DE RIKEN FULL-LENGTH ENRICHED LIBRARY, CLONE:2B10440D03, FULL INSERT  
DE SEQUENCE, full insert sequence.  
DE OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=99279253; PubMed=10349636;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=21085660; PubMed=11217851;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RL 60,770 full-length cDNAs.";  
RN Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20499374; PubMed=11042159;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RL prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20350913; PubMed=11076861;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-304-format  
RL sequencing pipeline with 384 multipipillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,



```

RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoch M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sojabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR HSSP: AK013274; BAB38762.1; -.
DR EMBL: P12367; 1R2A.
DR GO: GO:0008603; F:camp-dependent protein kinase regulator act. .; IEA.
DR GO: GO:0007165; P:signal transduction; IEA.
DR InterPro: IPR003117; RIIA.
DR Pfam: PF02197; RIIA; 1.
DR SMART: SM00394; RIIA; 1.
DR PROSITE: PS00430; TONB DEPENDENT REC 1; UNKNOWN 1.
SQ SEQUENCE 134 AA; 15229 MW; 19C837608E548A8_CRC64;

Query Match
Best Local Similarity 91.7%; Score 22; DB 2; Length 134;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 16 GYTVE 20

RESULT 54
Q853T4 ID Q853T4 PRELIMINARY; PRT; 135 AA.
AC Q853T4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Gp233.
GN Name=233;
OS Mycobacteriophage Omega.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=205879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22592860; PubMed=12705866;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.P.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003).
DR EMBL: AY129338; AAN12874.1; -.
SQ SEQUENCE 135 AA; 15098 MW; 9288CSAACFAF86BE_CRC64;

Query Match
Best Local Similarity 91.7%; Score 22; DB 2; Length 135;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 87 GYSVE 91

RESULT 55
RUVX_DEIRA ID RUVX_DEIRA STANDARD; PRT; 136 AA.
AC Q9RR12;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Putative Holliday junction resolvase (EC 3.1.-.-).
GN OrderedLocNames=DR2509;

OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -!- FUNCTION: Could be a nuclease that resolves Holliday junction intermediates in genetic recombination.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the yqgF HJR family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE002080; AAF12050.1; -.
DR PIR: D75265; D75265.
DR TIGR: DR2509; -.
DR HAMAP: MF_00651; -.
DR InterPro: IPR005227; Cons.hypoth250.
DR InterPro: IPR006641; YqgFC.
DR Pfam: PF03652; UFF0081; 1.
DR SMART: SM00732; YqgFc; 1.
DR TIGRFAMs: TIGR00250; Cons.hypoth250; 1.
KW Complete proteome; DNA recombination; DNA repair; Hydrolase; Nuclease.
SQ SEQUENCE 136 AA; 14819 MW; 7595D5353D2F21DE_CRC64;

Query Match
Best Local Similarity 91.7%; Score 22; DB 1; Length 136;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 98 GYTVE 102

RESULT 56
Q9HY15 ID Q9HY15 PRELIMINARY; PRT; 136 AA.
AC Q9HY15;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=PA3611;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,

```

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964 (2000).  
 DR EMBL; AE004781; AAC06999.1; --  
 DR FIR; H83194; H83194.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 136 AA; 14986 MW; 93DBDD79B601DD43 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 136;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 54 GYTVE 58

RESULT 57  
 YH92 ARCFU STANDARD; PRT; 137 AA.  
 ID YH92 ARCFU STANDARD; PRT; 137 AA.  
 AC Q28482;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Hypothetical protein AF1792;  
 GN OrderedLocusNames=AF1792;  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 OX NCBI\_TaxID=2234;  
 RN [1]

SEQUENCE FROM N.A.  
 STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RC MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Karlavage A.R., Graham D.E., Kyriades N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,  
 RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,  
 RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,  
 RA Usterback T.R., Cotton M.D., Spriggs T., Artiach P., Kaine B.P.,  
 RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,  
 RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,  
 RA Woese C.R., Venter J.C.;  
 RA "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon *Archaeoglobus fulgidus*.";  
 RL Nature 390:364-370 (1997).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----

DR EMBL; AE000979; AAC89473.1; --  
 DR TIGR; AF1792; --  
 KW Complete proteome; Hypothetical protein; Transmembrane.  
 FT TRANSMEM 20 39 Potential.  
 FT TRANSMEM 44 61 Potential.  
 FT TRANSMEM 86 105 Potential.  
 FT TRANSMEM 109 131 Potential.  
 SQ SEQUENCE 137 AA; 14779 MW; 5B6366C220ED337B CRC64;

Query Match 91.7%; Score 22; DB 1; Length 137;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 36 GYAVE 40

RESULT 58  
 Q9KHT5 PRELIMINARY; PRT; 139 AA.  
 ID Q9KHT5 PRELIMINARY; PRT; 139 AA.  
 AC Q9KHT5;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE Putative secreted protein.  
 OS Pseudomonas putida.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=20336420; PubMed=10877766;  
 RA Lee S.W., Cooksey D.A.;  
 RT "Genes expressed in *Pseudomonas putida* during colonization of a plant-  
 RT pathogenic fungus";  
 RL Appl. Environ. Microbiol. 66:2764-2772 (2000).  
 DR EMBL; AF249735; AAF64233.1; --  
 SQ SEQUENCE 139 AA; 15353 MW; 18DDE202D78BA3 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 139;  
 Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 ||||  
 Db 56 GYTVE 60

RESULT 59  
 Q88ME0 PRELIMINARY; PRT; 139 AA.  
 ID Q88ME0 PRELIMINARY; PRT; 139 AA.  
 AC Q88ME0;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=PP1633;  
 OS Pseudomonas putida (strain KT2440).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=160488;  
 RN [1]

SEQUENCE FROM N.A.  
 RP MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
 RA Brinkac L.M., Beanan M.J., DeBoer R.T., Daugherty S.C., Kolonay J.F.,  
 RA Hance J., Nelson W.C., White O., Peterson J.D., Khouri H.M.,  
 RA Madupu R., Chris Lee P., Holtzapfel E.K., Scanlan D., Tran K.,  
 RA Moazzaz A., Usterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,  
 RA Wedler H., Lauber J., Stjepandic D., Honeisel J., Straetz M., Heim S.,  
 RA Kiewitz C., Bisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,  
 RA Fraser C.M.;  
 RT "Complete genome sequence and comparative analysis of the  
 RT metabolically versatile *Pseudomonas putida* KT2440.";  
 RL Environ. Microbiol. 4:799-808 (2002).  
 DR EMBL; AE016780; AAN67254.1; --  
 DR TIGR; PP1633; --  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 139 AA; 15205 MW; 24A77CFE2A8B0A73 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 139;  
 Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 GYXVE 5
Db      56 GYTV 60

RESULT 60
Q919Z8
ID      Q9V9Z8      PRELIMINARY;      PRT;      142 AA.
AC      Q9V9Z8;
DT      01-NOV-1999 (TReMBLrel. 12, Created)
DT      01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE      Hypothetical protein APE2141.
GN      OrderedLocusNames=APE2141;
OS      Aeropyrum pernix.
OC      Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC      Desulfurococaceae; Aeropyrum.
OX      NCBI_TaxID=56636;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K1;
RX      MEDLINE=99310339; PubMed=10382966;
RA      Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA      Jin-no K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,
RA      Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA      Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo H. Y.,
RA      Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA      Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT      "Complete genome sequence of an aerobic hyper-thermophilic
RT      Crenarchaeon, Aeropyrum pernix K1."
RL      DNA Res. 6:83-101(1999).
DR      EMBL; AP000063; BAA81152.1; -.
DR      PIR; H72520; H72520.
KW      Complete proteome; Hypothetical protein.
SQ      SEQUENCE 142 AA; 15447 MW; 9479424B108DF13 CRC64;

Query Match      91.7%; Score 22; DB 2; Length 142;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      89 GYTV 93

RESULT 61
Q7S0W9
ID      Q7S0W9      PRELIMINARY;      PRT;      144 AA.
AC      Q7S0W9;
DT      01-MAR-2004 (TReMBLrel. 26, Created)
DT      01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT      01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE      Predicted protein.
GN      Name=NCU06955.1;
OS      Neurospora crassa.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC      Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX      NCBI_TaxID=5141;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=OR74A;
RA      Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA      Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA      Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA      Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA      Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA      Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA      Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA      Kamal M., Kamysseelis M., Mauceli E., Bielek C., Rudd S., Frishman D.,
RA      Kryzstofova S., Ramussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA      Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA      DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,

```

```

RA      Yarden O., Piamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA      Natvig D.O., Alex L.A., Mannheim G., Eboile D.J., Freitag M.,
RA      Paulsen I., Sachs M.S., Lander E.S., Nubbaum C., Birren B.;
RT      "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL      Nature 0:0-0(2003).
CC      -!- CAUTION: The sequence shown here is derived from an
CC      EMBL/GenBank/DDSI whole genome shotgun (WGS) entry which is
CC      preliminary data.
DR      EMBL; AABX0100474; EAA28999.1; -.
SQ      SEQUENCE 144 AA; 16047 MW; 8103DD3CC7B1ECE8 CRC64;

Query Match      91.7%; Score 22; DB 2; Length 144;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      30 GYAVE 34

RESULT 62
Q8BEZ2
ID      Q8BEZ2      PRELIMINARY;      PRT;      145 AA.
AC      Q8BEZ2;
DT      01-JUN-2003 (TReMBLrel. 24, Created)
DT      01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT      01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE      Transcriptional regulator, AsnC family.
GN      OrderedLocusNames=PP4308;
OS      Pseudomonas putida (strain KT2440).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Pseudomonas.
OX      NCBI_TaxID=160488;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22423060; PubMed=12534463;
RA      Nelson K.E., Weinell C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA      Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA      Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA      Haddad R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA      Manu C.I., Chris Lee P., Holtzapfel E.K., Scanlan D., Tran K.,
RA      Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA      Wedler H., Lauber J., Stjepandic D., Hobeisel J., Straetz M., Heim S.,
RA      Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmier B.,
RA      Fraser C.M.;
RT      "Complete genome sequence and comparative analysis of the
RT      metabolically versatile Pseudomonas putida KT2440."
RL      Environ. Microbiol. 4:799-808(2002).
CC      -!- SIMILARITY: Contains 1 HTH asnC-type DNA-binding domain.
DR      EMBL; AE016790; AAN69888.1; -.
DR      TIGR; PP4308; -.
DR      GO; GO:0005622; C:intracellular; IEA.
DR      GO; GO:0003700; F:transcription factor activity; IEA.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR      InterPro; IPR011008; Dimer_A_barrel.
DR      InterPro; IPR000485; HTH_AsnC_lrp.
DR      InterPro; IPR002197; HTH_Fis.
DR      Pfam; PF01037; AsnC_trans_reg; 1.
DR      PRINTS; PR00033; HTHASNC.
DR      PRINTS; PR01590; HTHFIS.
DR      SMART; SM00344; HTH_ASN_C; 1.
DR      TIGRFAMs; TIGR01199; HTH_fis; 1.
DR      PROSITE; PS50956; HTH_ASN_C_2; 1.
KW      Complete proteome; DNA-binding; Transcription regulation.
SQ      SEQUENCE 145 AA; 16411 MW; 6BD39C0657180806 CRC64;

Query Match      91.7%; Score 22; DB 2; Length 145;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      1 GYXVE 5

```

Poch O., Prieur D., Querellou J., Kipp R., Inlery J.-C., Van der Oost J., Weissbach J., Zivanovic Y., Forterre P.;  
"An integrated analysis of the genome of the hyperthermophilic archaeon Pyrococcus abyssi.";  
Mol. Microbiol. 47:1495-1512(2003).  
-!- FUNCTION: Molecular chaperone capable of stabilizing a range of proteins. Seems to fulfil an ATP-independent, HSP70-like function in archaeal de novo protein folding (By similarity).  
-!- SUBUNIT: Heterohexamer of two alpha and four beta subunits (By similarity).  
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
-!- SIMILARITY: Belongs to the prefoldin alpha subunit family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC ENBL; AJ248287; CAB50428.1; -.  
DR PIR; G75066; G75066.  
DR HAMAP; MF 00308; -. 1.  
DR InterPro; IPR004127; DUF232.  
DR InterPro; IPR009053; Prefoldin.  
DR Pfam; PF02996; Prefoldin; 1.  
DR Chaperone; Complete proteome.  
KW CHAPERONE; 148 AA; 16561 MW; DCE05593AF2EA1B8 CRC64;  
SQ SEQUENCE 148 AA; 16561 MW; DCE05593AF2EA1B8 CRC64;  
-----  
Query Match 91.7%; Score 22; DB 1; Length 148;  
Best Local Similarity 80.0%; Pred. No. 1.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
-----  
Qy 1 GYXVE 5  
|| ||  
Db 86 GYAVE 90  
-----  
RESULT 65  
PFDA PYRHO STANDARD; PRT; 148 AA.  
AC O58263;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Prefoldin alpha subunit (GimC alpha subunit).  
DE Name=pfda; OrderedLocNames=PH0327;  
GN Pyrococcus horikoshii.  
OS Archaea; Euryarchaeota; Thermococci; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=53953;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=OT3;  
RX MEDLINE=98344137; PubMed=9679194;  
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi Y., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:555-76(1998).  
CC -!- FUNCTION: Molecular chaperone capable of stabilizing a range of proteins. Seems to fulfil an ATP-independent, HSP70-like function in archaeal de novo protein folding (By similarity).  
CC -!- SUBUNIT: Heterohexamer of two alpha and four beta subunits (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the prefoldin alpha subunit family.  
-----

Db 50 GYTV 54

```
RESULT 63
PFDA_PYRAB
ID -PFDA PYRFU STANDARD; PRT; 146 AA.
AC QBU3T0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Prefoldin alpha subunit (GimC alpha subunit).
GN Name=pfda; OrderedLocusNames=PF0375;
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
NCBI_TaxID=2261;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Welts R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Molecular chaperone capable of stabilizing a range of
CC proteins. Seems to fulfil an ATP-independent, HSP70-like function
CC in archaeal de novo protein folding (By similarity).
CC -!- SUBUNIT: Heterohexamer of two alpha and four beta subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the prefoldin alpha subunit family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AE010161; AAL80499.1; -.
DR HAMAP; MF 00308; -; 1.
DR InterPro; IPR004127; DUF232.
DR InterPro; IPR009053; Prefoldin.
DR Pfam; PF02996; Prefoldin; 1.
KW Chaperone; Complete proteome.
SQ SEQUENCE 146 AA; 16509 MW; B736A8AFD482DD96 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 146;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 85 GYAVE 89

RESULT 64
PFDA_PYRAB
ID PFDA PYRAB STANDARD; PRT; 148 AA.
AC Q9UYI4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Prefoldin alpha subunit (GimC alpha subunit).
GN Name=pfda; OrderedLocusNames=PYRAB15230; ORFNames=PAB1006;
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
NCBI_TaxID=29292;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RA MEDLINE=22511545; PubMed=12622808;
RX Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; AP000002; BAA29615.1; ALT\_INIT.

CC DR HAMAP; MF 00308; -; 1.

CC DR InterPro; IPR004127; DUF232.

CC DR InterPro; IPR009053; Prefoldin.

CC DR Pfam; PF02996; Prefoldin; 1.

KW Chapterone; Complete proteome.

SQ SEQUENCE 148 AA; 16603 MW; 46AB8EA20A4AC8BD CRC64;

Query Match 91.7%; Score 22; DB 1; Length 148;  
Best Local Similarity 80.0%; Pred. No. 1.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVE 5

|||

86 GYAVE 90

RESULT 66

RL9 LISIN

ID RL9 LISIN STANDARD; PRT; 148 AA.

AC Q92FQ7;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE 50S ribosomal protein L9.

GN Name=rplI; OrderedLocusNames=lin0046;

OS Listeria innocua.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OX NCBI\_TaxID=1642;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=CLIP 11262 / Serovar 6a;

EX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;

RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,

RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,

RT "Comparative genomics of Listeria species.";

RL Science 294:849-852(2001).

CC -!- FUNCTION: Binds to the 23S rRNA (By similarity).

CC -!- SIMILARITY: Belongs to the L9p family of ribosomal proteins.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; AL596163; CAC95279.1; -.

CC DR PIR; AG1438; AG1438.

CC DR HSP; P02417; 1CQU.

CC DR ListiList; LIN00046; -.

CC DR HAMAP; MF 00503; -; 1.

CC DR InterPro; IPR009027; L9\_N\_like.

CC DR InterPro; IPR000244; Ribosomal L9.

CC Pfam; PF03948; Ribosomal L9 C; 1.

CC

DR Pfam; PF01281; Ribosomal L9\_N; 1.  
DR TIGRFAMs; TIGR00158; L9; 1.  
DR PROSITE; PS00651; RIBOSOMAL\_L9; 1.  
KW Complete proteome; Ribosomal protein; rRNA-binding.  
SQ SEQUENCE 148 AA; 16137 MW; 6D17B9E9B68A7755 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 148;  
Best Local Similarity 80.0%; Pred. No. 1.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVE 5

|||

34 GYAVE 38

RESULT 67

RL9 LISMO

ID RL9 LISMO STANDARD; PRT; 148 AA.

AC Q8V8R2;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE 50S ribosomal protein L9.

GN Name=rplI; OrderedLocusNames=lmo0053;

OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OX NCBI\_TaxID=1639;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=EGD-e / Serovar 1/2a;

EX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;

RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,

RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,

RT "Comparative genomics of Listeria species.";

RL Science 294:849-852(2001).

CC -!- FUNCTION: Binds to the 23S rRNA (By similarity).

CC -!- SIMILARITY: Belongs to the L9p family of ribosomal proteins.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; AL591973; CAC98268.1; -.

CC DR PIR; AF1081; AF1081.

CC DR HSP; P02417; 1CQU.

CC DR ListiList; LMO00053; -.

CC DR HAMAP; MF\_00503; -; 1.

CC DR InterPro; IPR009027; L9\_N\_like.

CC DR InterPro; IPR000244; Ribosomal L9.

CC Pfam; PF03948; Ribosomal L9 C; 1.

CC Pfam; PF01281; Ribosomal L9\_N; 1.

CC TIGRFAMs; TIGR00158; L9; 1.

CC PROSITE; PS00651; RIBOSOMAL\_L9; 1.

KW Complete proteome; Ribosomal protein; rRNA-binding.

SQ SEQUENCE 148 AA; 16164 MW; 74E94FF5CC1F9336 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 148;  
Best Local Similarity 80.0%; Pred. No. 1.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

[illegible][illegible]

```
SQ SEQUENCE 151 AA; 17566 MW; 27B09295FDE84814 CRC64;
Query Match 91.7%; Score 22; DB 2; Length 151;
Best Local Similarity 80.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 106 GYSVE 110

RESULT 72
AAR10433 PRELIMINARY; PRT; 151 AA.
AC AAR10433;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Enterococcus faecium (Streptococcus faecium).
OG Plasmid pppM1000.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1352;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=664.1H1;
RA Davis I.J., Roberts A.P., Mullany P.;
RT "Novel Genetic Organization of a Transferable Mercury Resistance
RT Operon from Enterococcus faecium.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY351675; AAR10433.1; -.
KW Hypothetical protein; Plasmid.
FT NON TER 1
SQ SEQUENCE 151 AA; 17566 MW; 27B09295FDE84814 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 151;
Best Local Similarity 80.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 106 GYSVE 110

RESULT 73
YE83 METAJ STANDARD; PRT; 152 AA.
AC Q59878;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical protein MJ1483.
GN OrderedLocNames=MJ1483;
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Scott J.L., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Utterback T.R., Kelley J.M., Weidman J.P., Fuhrmann J.L., Nguyen D.,
RA Cotton M.D., Roberts K.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Klenk H.-P., Fraser C.M., Hirst M.A., Kaine B.P., Borodovsky M.,
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RC Science 273:1058-1073 (1996).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U67589; AAB99498.1; -.
DR PIR; B64485; B64485.
DR TIGR; MJ1483; -.
KW Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 7 27
SQ SEQUENCE 152 AA; 17659 MW; 90227132448B9802 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 152;
Best Local Similarity 80.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 107 GYSVE 111

RESULT 74
Q7WOH6 PRELIMINARY; PRT; 153 AA.
AC Q7WOH6;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative exported protein.
GN OrderedLocNames=BP0156;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640411; CAE40535.1; -.
DR InterPro; IPR007332; DUF411.
DR Pfam; PF04214; DUF411; 1.
KW Complete proteome.
SQ SEQUENCE 153 AA; 16516 MW; 78DC41BB19F593D3 CRC64;

Query Match 91.7%; Score 22; DB 2; Length 153;
Best Local Similarity 80.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
DB 86 GYTVE 90

RESULT 75
Q7W3N1 PRELIMINARY; PRT; 153 AA.
AC Q7W3N1;
ID Q7W3N1;
RT jannaschii.";
```

Thu Nov 4 17:32:36 2004

01-OCT-2003 (TrEMBLrel. 25, Created)  
 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative exported protein.  
 GN Ordered locus names=BPP4000;  
 OS Bordetella parapertussis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=12822 / ATCC BAA-587;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi227;  
 RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,  
 Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 Cardeno-Taraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 Feltwell T., Gobie A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,  
 Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL; BX640435; CAE39283.1; -.  
 DR InterPro; IPR007332; DUF411.  
 DR Pfam; PF04214; DUF411; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 153 AA; 16488 MW; 63DC41BB19F593CC CRC64;

Query Match 91.7%; Score 22; DB 2; Length 153;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
 Db 86 GYTV 90

Search completed: November 1, 2004, 21:35:01  
 Job time : 161.5 secs





101	22	91.7	885	4	US-09-538-092-1319	Sequence 1319, Ap	174	21	87.5	197	3	US-09-813-918-4	Sequence 4, Appli
102	22	91.7	1237	1	US-08-241-853-2	Sequence 2, Appli	175	21	87.5	197	4	US-10-060-311-4	Sequence 5, Appli
103	22	91.7	1237	2	US-08-850-917-2	Sequence 2, Appli	176	21	87.5	201	1	US-07-929-580B-5	Sequence 5, Appli
104	22	91.7	1395	3	US-09-540-245A-15	Sequence 15, Appl	177	21	87.5	201	4	US-09-198-452A-548	Sequence 548, App
105	22	91.7	1572	4	US-09-710-279-2906	Sequence 2906, Ap	178	21	87.5	205	4	US-09-710-279-3020	Sequence 3020, Ap
106	22	91.7	3892	4	US-09-328-352-5503	Sequence 5503, Ap	179	21	87.5	206	3	US-09-041-889-6	Sequence 6, Appli
107	21	87.5	12	1	US-08-036-555B-164	Sequence 164, App	180	21	87.5	206	3	US-08-837-058-6	Sequence 6, Appli
108	21	87.5	12	1	US-08-469-569-164	Sequence 164, App	181	21	87.5	206	4	US-09-417-264-6	Sequence 6, Appli
109	21	87.5	12	1	US-08-249-322A-164	Sequence 164, App	182	21	87.5	206	4	US-09-107-532A-6492	Sequence 6492, Ap
110	21	87.5	12	1	US-08-469-526A-164	Sequence 164, App	183	21	87.5	208	1	US-07-935-309-2	Sequence 2, Appli
111	21	87.5	12	2	US-08-734-591A-164	Sequence 164, App	184	21	87.5	208	2	US-08-039-364-2	Sequence 2, Appli
112	21	87.5	12	2	US-08-469-660-164	Sequence 164, App	185	21	87.5	208	3	US-08-718-904-5	Sequence 5, Appli
113	21	87.5	12	3	US-08-470-335-164	Sequence 164, App	186	21	87.5	208	3	US-08-718-904-7	Sequence 7, Appli
114	21	87.5	12	3	US-08-735-021-164	Sequence 164, App	187	21	87.5	208	3	US-09-181-974-2	Sequence 2, Appli
115	21	87.5	12	3	US-08-734-664A-164	Sequence 164, App	188	21	87.5	208	3	US-09-158-710-2	Sequence 2, Appli
116	21	87.5	12	3	US-08-470-339-164	Sequence 164, App	189	21	87.5	208	3	US-09-134-001C-4132	Sequence 4132, Ap
117	21	87.5	12	4	US-08-467-602-164	Sequence 164, App	190	21	87.5	208	4	US-09-518-950-2	Sequence 2, Appli
118	21	87.5	12	5	PCT-US94-05083C-160	Sequence 160, App	191	21	87.5	208	4	US-09-449-249-5	Sequence 5, Appli
119	21	87.5	12	5	PCT-US95-06846A-164	Sequence 164, App	192	21	87.5	208	4	US-09-449-249-7	Sequence 7, Appli
120	21	87.5	16	4	US-09-775-743C-15	Sequence 15, Appl	193	21	87.5	208	4	US-10-138-158-18	Sequence 18, Appl
121	21	87.5	51	3	US-09-399-494-16	Sequence 16, Appl	194	21	87.5	208	4	US-09-248-796A-18276	Sequence 18276, A
122	21	87.5	21	3	US-09-399-494-17	Sequence 17, Appl	195	21	87.5	211	4	US-09-248-796A-18236	Sequence 18236, A
123	21	87.5	21	3	US-09-946-180B-5	Sequence 5, Appli	196	21	87.5	212	3	US-09-041-889-1	Sequence 1, Appli
124	21	87.5	32	4	US-09-230-041-11	Sequence 11, Appl	197	21	87.5	212	3	US-08-837-058-1	Sequence 1, Appli
125	21	87.5	51	3	US-07-946-180B-8	Sequence 8, Appli	198	21	87.5	212	4	US-09-417-264-1	Sequence 1, Appli
126	21	87.5	53	1	US-08-313-050-18	Sequence 18, Appl	199	21	87.5	212	4	US-08-486-049-4	Sequence 4, Appli
127	21	87.5	62	3	US-07-946-180B-9	Sequence 9, Appli	200	21	87.5	213	4	US-09-252-991A-23391	Sequence 23391, A
128	21	87.5	63	4	US-09-270-767-57725	Sequence 57725, A	201	21	87.5	215	3	US-08-914-375C-29	Sequence 29, Appl
129	21	87.5	85	4	US-09-248-796A-23917	Sequence 23917, A	202	21	87.5	215	3	US-09-328-352-4289	Sequence 4289, Ap
130	21	87.5	95	3	US-09-374-135-8	Sequence 8, Appli	203	21	87.5	218	3	US-09-041-889-4	Sequence 4, Appli
131	21	87.5	102	2	US-08-808-982-8	Sequence 8, Appli	204	21	87.5	218	3	US-08-837-058-4	Sequence 4, Appli
132	21	87.5	103	3	US-09-306-902A-8	Sequence 8, Appli	205	21	87.5	218	4	US-09-417-264-4	Sequence 4, Appli
133	21	87.5	103	3	US-09-041-889-39	Sequence 39, Appl	206	21	87.5	220	3	US-09-041-889-2	Sequence 2, Appli
134	21	87.5	103	4	US-09-417-264-39	Sequence 39, Appl	207	21	87.5	220	3	US-08-837-058-2	Sequence 2, Appli
135	21	87.5	103	4	US-09-732-210-763	Sequence 763, App	208	21	87.5	220	4	US-09-417-264-2	Sequence 2, Appli
136	21	87.5	105	4	US-09-513-999C-5473	Sequence 5473, Ap	209	21	87.5	222	3	US-09-041-889-3	Sequence 3, Appli
137	21	87.5	107	4	US-09-107-532A-5829	Sequence 5829, Ap	210	21	87.5	222	3	US-08-837-058-3	Sequence 3, Appli
138	21	87.5	109	4	US-09-134-000C-5763	Sequence 5763, Ap	211	21	87.5	222	4	US-09-417-264-3	Sequence 3, Appli
139	21	87.5	111	4	US-09-270-767-58884	Sequence 58884, A	212	21	87.5	223	1	US-07-708-885B-3	Sequence 3, Appli
140	21	87.5	115	4	US-09-173-300-13	Sequence 13, Appl	213	21	87.5	223	1	US-07-714-386-3	Sequence 3, Appli
141	21	87.5	116	3	US-09-041-889-38	Sequence 38, Appl	214	21	87.5	223	3	US-07-708-888A-3	Sequence 3, Appli
142	21	87.5	116	4	US-09-417-264-38	Sequence 38, Appl	215	21	87.5	224	3	US-08-963-901-6	Sequence 6, Appli
143	21	87.5	118	4	US-09-585-173B-20	Sequence 20, Appl	216	21	87.5	224	4	US-09-710-279-496	Sequence 496, App
144	21	87.5	119	4	US-09-270-767-58586	Sequence 58586, A	217	21	87.5	224	4	US-09-710-279-1366	Sequence 1366, Ap
145	21	87.5	120	4	US-09-270-767-32195	Sequence 32195, A	218	21	87.5	225	4	US-09-107-532A-6303	Sequence 6303, Ap
146	21	87.5	120	4	US-09-270-767-47412	Sequence 47412, A	219	21	87.5	225	4	US-09-041-889-32	Sequence 32, Appl
147	21	87.5	123	3	US-09-134-001C-3259	Sequence 3259, Ap	220	21	87.5	226	4	US-09-417-264-32	Sequence 32, Appl
148	21	87.5	126	4	US-09-583-110-5175	Sequence 5175, Ap	221	21	87.5	228	3	US-09-134-001C-4373	Sequence 4373, Ap
149	21	87.5	128	4	US-09-134-000C-3570	Sequence 3570, Ap	222	21	87.5	228	3	US-09-134-001C-5384	Sequence 5384, Ap
150	21	87.5	131	4	US-09-328-352-4505	Sequence 4505, Ap	223	21	87.5	229	3	US-09-078-691-4	Sequence 4, Appli
151	21	87.5	135	4	US-09-134-000C-4619	Sequence 4619, Ap	224	21	87.5	230	4	US-09-248-796A-16642	Sequence 16642, A
152	21	87.5	139	4	US-09-710-279-3044	Sequence 3044, Ap	225	21	87.5	231	4	US-09-583-110-4922	Sequence 4922, Ap
153	21	87.5	139	4	US-09-710-279-3302	Sequence 3302, Ap	226	21	87.5	233	2	US-08-771-783-2	Sequence 2, Appli
154	21	87.5	143	3	US-08-858-207A-441	Sequence 441, App	227	21	87.5	236	4	US-09-107-532A-6873	Sequence 6873, Ap
155	21	87.5	145	4	US-09-107-532A-3705	Sequence 3705, Ap	228	21	87.5	237	4	US-09-489-039A-7234	Sequence 7234, Ap
156	21	87.5	154	4	US-09-538-092-1383	Sequence 1383, Ap	229	21	87.5	238	4	US-09-107-532A-6268	Sequence 6268, Ap
157	21	87.5	154	4	US-09-513-999C-8054	Sequence 8054, Ap	230	21	87.5	239	4	US-09-270-767-42453	Sequence 42453, A
158	21	87.5	155	3	US-09-203-716-4	Sequence 4, Appli	231	21	87.5	243	4	US-09-107-532A-4665	Sequence 4665, Ap
159	21	87.5	155	3	US-09-684-254-4	Sequence 4, Appli	232	21	87.5	245	3	US-09-134-001C-3779	Sequence 3779, Ap
160	21	87.5	155	4	US-09-409-926-5	Sequence 5, Appli	233	21	87.5	246	4	US-09-252-991A-23985	Sequence 23985, A
161	21	87.5	157	4	US-09-270-767-36537	Sequence 36537, A	234	21	87.5	250	4	US-09-710-279-1686	Sequence 1686, Ap
162	21	87.5	157	4	US-09-270-767-71574	Sequence 51754, A	235	21	87.5	253	4	US-09-543-681A-5796	Sequence 5796, Ap
163	21	87.5	158	3	US-09-041-889-40	Sequence 40, Appl	236	21	87.5	254	4	US-09-583-110-5019	Sequence 5019, Ap
164	21	87.5	158	4	US-09-417-264-40	Sequence 40, Appl	237	21	87.5	259	4	US-09-252-991A-32989	Sequence 32989, A
165	21	87.5	161	4	US-09-107-532A-6820	Sequence 6820, Ap	238	21	87.5	259	4	US-09-134-000C-4642	Sequence 4642, Ap
166	21	87.5	163	4	US-09-538-092-600	Sequence 600, App	239	21	87.5	260	4	US-09-585-173B-22	Sequence 22, Appl
167	21	87.5	165	4	US-09-489-039A-14030	Sequence 14030, A	240	21	87.5	264	4	US-09-328-352-6490	Sequence 6490, Ap
168	21	87.5	168	4	US-09-252-991A-25254	Sequence 25254, A	241	21	87.5	264	4	US-09-886-319A-67	Sequence 67, Appl
169	21	87.5	172	4	US-09-248-796A-14431	Sequence 14431, A	242	21	87.5	265	4	US-09-886-319A-68	Sequence 68, Appl
170	21	87.5	173	4	US-09-583-110-4353	Sequence 4353, Ap	243	21	87.5	265	4	US-09-252-991A-29625	Sequence 29625, A
171	21	87.5	180	4	US-09-252-991A-26385	Sequence 26385, A	244	21	87.5	269	4	US-09-634-238-222	Sequence 222, App
172	21	87.5	181	4	US-09-543-681A-5968	Sequence 5968, Ap	245	21	87.5	270	4	US-09-230-041-2	Sequence 2, Appli
173	21	87.5	192	4	US-09-107-532A-6267	Sequence 6267, Ap	246	21	87.5	276	4	US-09-328-352-6348	Sequence 6348, Ap

247	21	87.5	279	4	US-09-543-681A-7858	Sequence 7858, Ap	320	21	87.5	358	4	US-09-710-279-18	Sequence 18, Appl
248	21	87.5	281	1	US-08-487-748A-9	Sequence 9, Appl	321	21	87.5	359	4	US-09-538-092-111	Sequence 11, App
249	21	87.5	281	1	US-08-487-748A-10	Sequence 10, Appl	322	21	87.5	361	3	US-09-120-365-75	Sequence 75, Appl
250	21	87.5	281	3	US-08-480-070C-10	Sequence 10, Appl	323	21	87.5	361	3	US-09-515-039-75	Sequence 75, Appl
251	21	87.5	281	3	US-08-828-525-10	Sequence 10, Appl	324	21	87.5	361	3	US-09-134-001C-3862	Sequence 3862, Ap
252	21	87.5	281	3	US-08-609-583A-10	Sequence 10, Appl	325	21	87.5	361	4	US-09-724-566A-75	Sequence 75, Appl
253	21	87.5	281	3	US-08-937-399-10	Sequence 10, Appl	326	21	87.5	362	1	US-08-118-270-15	Sequence 15, Appl
254	21	87.5	281	4	US-09-310-367-10	Sequence 10, Appl	327	21	87.5	362	3	US-09-625-188-8	Sequence 8, Appl
255	21	87.5	281	4	US-09-032-337-10	Sequence 10, Appl	328	21	87.5	362	5	PCT-US93-08528-15	Sequence 5403, Ap
256	21	87.5	281	4	US-09-464-231-10	Sequence 10, Appl	329	21	87.5	362	5	PCT-US93-08528-15	Sequence 15, Appl
257	21	87.5	282	4	US-09-252-991A-26679	Sequence 26679, A	330	21	87.5	364	4	US-09-710-279-3076	Sequence 3076, Ap
258	21	87.5	283	3	US-08-807-151-1	Sequence 1, Appl	331	21	87.5	367	4	US-09-134-000C-6173	Sequence 6173, Ap
259	21	87.5	283	3	US-09-478-957-1	Sequence 1, Appl	332	21	87.5	369	4	US-09-543-681A-7314	Sequence 7314, Ap
260	21	87.5	286	4	US-09-134-000C-4899	Sequence 4899, Ap	333	21	87.5	369	4	US-09-134-001C-4798	Sequence 4798, Ap
261	21	87.5	288	3	US-09-273-839A-8	Sequence 8, Appl	334	21	87.5	373	4	US-09-248-796A-18227	Sequence 18227, A
262	21	87.5	290	4	US-09-252-991A-32494	Sequence 32494, A	335	21	87.5	374	4	US-09-543-681A-5310	Sequence 5310, Ap
263	21	87.5	293	4	US-09-248-796A-18524	Sequence 18524, A	336	21	87.5	374	4	US-09-724-566A-71	Sequence 71, Appl
264	21	87.5	294	4	US-09-252-991A-26953	Sequence 26953, A	337	21	87.5	378	4	US-09-325-932A-158	Sequence 158, App
265	21	87.5	296	4	US-09-583-110-2963	Sequence 2963, A	338	21	87.5	378	4	US-09-489-039A-10489	Sequence 10489, A
266	21	87.5	297	4	US-09-107-532A-5306	Sequence 5306, Ap	339	21	87.5	378	4	US-09-270-767-59898	Sequence 59898, A
267	21	87.5	301	4	US-09-583-110-3523	Sequence 3523, Ap	340	21	87.5	379	4	US-09-252-991A-31570	Sequence 31570, A
268	21	87.5	302	3	US-09-282-305-14	Sequence 14, Appl	341	21	87.5	381	4	US-09-673-395A-363	Sequence 363, App
269	21	87.5	302	4	US-09-883-720-14	Sequence 14, Appl	342	21	87.5	383	4	US-08-311-731A-193	Sequence 193, App
270	21	87.5	303	2	US-08-599-171A-32	Sequence 32, Appl	343	21	87.5	385	4	US-09-328-352-4488	Sequence 4488, Ap
271	21	87.5	303	2	US-08-646-580B-32	Sequence 32, Appl	344	21	87.5	388	4	US-09-489-039A-7745	Sequence 7745, Ap
272	21	87.5	303	3	US-09-069-226-32	Sequence 32, Appl	345	21	87.5	390	4	US-09-328-352-4916	Sequence 4916, Ap
273	21	87.5	303	3	US-09-412-184-32	Sequence 32, Appl	346	21	87.5	390	4	US-09-724-566A-70	Sequence 70, Appl
274	21	87.5	305	3	US-09-282-305-12	Sequence 12, Appl	347	21	87.5	392	4	US-09-489-039A-11256	Sequence 11256, A
275	21	87.5	305	4	US-09-883-720-12	Sequence 12, Appl	348	21	87.5	392	4	US-09-540-236-3310	Sequence 3310, Ap
276	21	87.5	305	4	US-09-868-061-9	Sequence 9, Appl	349	21	87.5	395	4	US-09-724-566A-68	Sequence 68, Appl
277	21	87.5	305	4	US-09-452-937A-44	Sequence 44, Appl	350	21	87.5	405	4	US-09-252-991A-29407	Sequence 29407, A
278	21	87.5	307	4	US-09-173-300-15	Sequence 15, Appl	351	21	87.5	406	4	US-09-252-991A-22867	Sequence 22867, A
279	21	87.5	311	3	US-09-282-305-16	Sequence 16, Appl	352	21	87.5	407	4	US-09-724-566A-58	Sequence 58, Appl
280	21	87.5	311	4	US-09-883-720-16	Sequence 16, Appl	353	21	87.5	408	4	US-09-248-796A-15721	Sequence 15721, A
281	21	87.5	311	4	US-09-252-991A-19489	Sequence 19489, A	354	21	87.5	409	4	US-09-543-681A-5451	Sequence 5451, Ap
282	21	87.5	312	4	US-09-540-236-2460	Sequence 2460, Ap	355	21	87.5	412	1	US-07-712-833A-3	Sequence 3, Appl
283	21	87.5	313	3	US-09-134-001C-5185	Sequence 5185, Ap	356	21	87.5	412	4	US-09-543-681A-7895	Sequence 7895, Ap
284	21	87.5	315	1	US-07-729-099-1	Sequence 1, Appl	357	21	87.5	413	1	US-08-444-734A-5	Sequence 5, Appl
285	21	87.5	315	1	US-08-257-392-1	Sequence 1, Appl	358	21	87.5	413	1	US-08-087-772A-17	Sequence 17, Appl
286	21	87.5	315	3	US-08-770-035-1	Sequence 1, Appl	359	21	87.5	413	2	US-08-467-568-12	Sequence 12, Appl
287	21	87.5	316	3	US-09-078-691-2	Sequence 2, Appl	360	21	87.5	413	2	US-09-030-582-12	Sequence 12, Appl
288	21	87.5	319	4	US-08-311-731A-63	Sequence 63, Appl	361	21	87.5	413	4	US-09-811-286-2	Sequence 2, Appl
289	21	87.5	320	3	US-08-852-730-9	Sequence 9, Appl	362	21	87.5	413	4	US-09-134-000C-6481	Sequence 6481, Ap
290	21	87.5	320	4	US-09-248-796A-17061	Sequence 17061, A	363	21	87.5	413	4	US-09-936-499-1	Sequence 1, Appl
291	21	87.5	323	1	US-08-404-445-3	Sequence 3, Appl	364	21	87.5	413	5	PCT-US91-00909-4	Sequence 2, Appl
292	21	87.5	324	4	US-09-328-352-4636	Sequence 4636, Ap	365	21	87.5	413	5	PCT-US91-00909-4	Sequence 4, Appl
293	21	87.5	324	4	US-09-270-767-42432	Sequence 42432, A	366	21	87.5	419	4	US-09-724-566A-57	Sequence 57, Appl
294	21	87.5	326	3	US-09-564-805-231	Sequence 231, App	367	21	87.5	419	4	US-09-270-767-45469	Sequence 45469, A
295	21	87.5	327	4	US-10-067-443-5	Sequence 5, Appl	368	21	87.5	420	4	US-09-328-352-7296	Sequence 7296, Ap
296	21	87.5	333	4	US-09-252-991A-29336	Sequence 29336, A	369	21	87.5	422	4	US-09-724-566A-60	Sequence 60, Appl
297	21	87.5	334	3	US-08-975-762-28	Sequence 28, Appl	370	21	87.5	422	4	US-09-252-991A-24886	Sequence 24886, A
298	21	87.5	334	3	US-08-821-324-28	Sequence 28, Appl	371	21	87.5	423	4	US-09-341-446B-9	Sequence 9, Appl
299	21	87.5	334	3	US-09-295-028-28	Sequence 28, Appl	372	21	87.5	425	3	US-09-305-984-22	Sequence 22, Appl
300	21	87.5	334	3	US-09-108-582-28	Sequence 28, Appl	373	21	87.5	425	3	US-09-305-984-24	Sequence 24, Appl
301	21	87.5	334	4	US-09-159-469-28	Sequence 28, Appl	374	21	87.5	425	4	US-09-548-372D-28	Sequence 28, Appl
302	21	87.5	334	4	US-09-693-542-28	Sequence 28, Appl	375	21	87.5	425	4	US-09-548-367D-28	Sequence 28, Appl
303	21	87.5	337	4	US-09-252-991A-24654	Sequence 24654, A	376	21	87.5	425	4	US-09-073-541A-24	Sequence 24, Appl
304	21	87.5	339	4	US-09-328-352-4674	Sequence 4674, Ap	377	21	87.5	425	4	US-09-073-541A-24	Sequence 22, Appl
305	21	87.5	339	4	US-09-107-532A-5514	Sequence 5514, Ap	378	21	87.5	425	4	US-09-551-853D-28	Sequence 28, Appl
306	21	87.5	341	4	US-09-270-767-44982	Sequence 44982, A	379	21	87.5	425	4	US-09-433-940-22	Sequence 22, Appl
307	21	87.5	342	4	US-09-489-039A-13595	Sequence 13595, A	380	21	87.5	425	4	US-09-493-940-24	Sequence 24, Appl
308	21	87.5	342	4	US-09-540-236-2385	Sequence 2385, Ap	381	21	87.5	425	4	US-09-416-901B-28	Sequence 28, Appl
309	21	87.5	344	3	US-09-199-637A-108	Sequence 108, App	382	21	87.5	425	4	US-09-583-110-3691	Sequence 2691, Ap
310	21	87.5	344	4	US-09-540-236-2452	Sequence 2452, Ap	383	21	87.5	425	4	US-09-583-110-5321	Sequence 5321, Ap
311	21	87.5	345	1	US-08-183-214-12	Sequence 12, Appl	384	21	87.5	425	4	US-09-548-376D-28	Sequence 28, Appl
312	21	87.5	348	3	US-08-855-910-8	Sequence 8, Appl	385	21	87.5	425	4	US-09-548-376D-28	Sequence 28, Appl
313	21	87.5	352	1	US-08-785-052-2	Sequence 2, Appl	386	21	87.5	425	4	US-09-548-373D-28	Sequence 28, Appl
314	21	87.5	352	2	US-08-913-581-2	Sequence 2, Appl	387	21	87.5	425	4	US-09-795-847B-28	Sequence 28, Appl
315	21	87.5	353	4	US-09-107-532A-4936	Sequence 4936, Ap	388	21	87.5	425	4	US-09-485-529-6	Sequence 6, Appl
316	21	87.5	353	4	US-09-252-991A-32769	Sequence 32769, A	389	21	87.5	425	4	US-09-869-414-28	Sequence 28, Appl
317	21	87.5	356	4	US-09-328-352-6258	Sequence 6258, Ap	390	21	87.5	425	4	US-09-548-366F-28	Sequence 28, Appl
318	21	87.5	357	3	US-09-134-001C-4405	Sequence 4405, Ap	391	21	87.5	427	4	US-09-107-532A-5530	Sequence 5530, Ap
319	21	87.5	358	4	US-09-252-991A-20554	Sequence 20554, A	392	21	87.5	428	4	US-09-548-372D-51	Sequence 51, Appl

393	21	87.5	428	4	US-09-548-367D-51	Sequence 51, Appl	466	21	87.5	459	4	US-09-794-927A-32	Sequence 32, Appl
394	21	87.5	428	4	US-09-551-853D-51	Sequence 51, Appl	467	21	87.5	459	4	US-09-548-373D-24	Sequence 24, Appl
395	21	87.5	428	4	US-09-416-901B-51	Sequence 51, Appl	468	21	87.5	459	4	US-09-548-373D-32	Sequence 32, Appl
396	21	87.5	428	4	US-09-548-376D-51	Sequence 51, Appl	469	21	87.5	459	4	US-09-795-847B-24	Sequence 24, Appl
397	21	87.5	428	4	US-09-794-927A-51	Sequence 51, Appl	470	21	87.5	459	4	US-09-795-847B-32	Sequence 32, Appl
398	21	87.5	428	4	US-09-548-373D-51	Sequence 51, Appl	471	21	87.5	459	4	US-09-869-414-24	Sequence 24, Appl
399	21	87.5	428	4	US-09-795-847B-51	Sequence 51, Appl	472	21	87.5	459	4	US-09-869-414-32	Sequence 32, Appl
400	21	87.5	428	4	US-09-869-414-51	Sequence 51, Appl	473	21	87.5	459	4	US-09-548-366F-24	Sequence 24, Appl
401	21	87.5	428	4	US-09-548-366F-51	Sequence 51, Appl	474	21	87.5	459	4	US-09-548-366F-32	Sequence 32, Appl
402	21	87.5	431	4	US-09-724-566A-74	Sequence 74, Appl	475	21	87.5	460	3	US-09-134-001C-5322	Sequence 5322, Ap
403	21	87.5	433	4	US-09-548-372D-26	Sequence 26, Appl	476	21	87.5	466	1	US-08-785-066-2	Sequence 2, Appl
404	21	87.5	433	4	US-09-548-367D-26	Sequence 26, Appl	477	21	87.5	466	3	US-09-007-355-2	Sequence 2, Appl
405	21	87.5	433	4	US-09-551-853D-26	Sequence 26, Appl	478	21	87.5	466	3	US-08-913-483-2	Sequence 2, Appl
406	21	87.5	433	4	US-09-416-901B-26	Sequence 26, Appl	479	21	87.5	467	4	US-09-538-092-409	Sequence 409, App
407	21	87.5	433	4	US-09-548-376D-26	Sequence 26, Appl	480	21	87.5	469	4	US-09-538-092-244	Sequence 244, App
408	21	87.5	433	4	US-09-794-927A-26	Sequence 26, Appl	481	21	87.5	469	4	US-09-538-092-416	Sequence 416, App
409	21	87.5	433	4	US-09-548-373D-26	Sequence 26, Appl	482	21	87.5	473	4	US-09-252-991A-29434	Sequence 29434, A
410	21	87.5	433	4	US-09-795-847B-26	Sequence 26, Appl	483	21	87.5	474	4	US-09-540-236-3271	Sequence 3271, Ap
411	21	87.5	433	4	US-09-869-414-26	Sequence 26, Appl	484	21	87.5	476	1	US-08-216-276A-33	Sequence 33, Appl
412	21	87.5	433	4	US-09-548-366F-26	Sequence 26, Appl	485	21	87.5	476	4	US-09-548-372D-6	Sequence 6, Appl
413	21	87.5	434	4	US-09-548-373D-53	Sequence 53, Appl	486	21	87.5	476	4	US-09-548-372D-73	Sequence 73, Appl
414	21	87.5	434	4	US-09-548-367D-53	Sequence 53, Appl	487	21	87.5	476	4	US-09-548-367D-6	Sequence 6, Appl
421	21	87.5	434	4	US-09-795-847B-53	Sequence 53, Appl	488	21	87.5	476	4	US-09-548-367D-73	Sequence 73, Appl
422	21	87.5	434	4	US-09-551-853D-53	Sequence 53, Appl	489	21	87.5	476	4	US-09-551-853D-6	Sequence 6, Appl
423	21	87.5	434	4	US-09-416-901B-53	Sequence 53, Appl	490	21	87.5	476	4	US-09-551-853D-73	Sequence 73, Appl
424	21	87.5	434	4	US-09-710-279-1456	Sequence 1456, Ap	491	21	87.5	476	4	US-09-416-901B-6	Sequence 6, Appl
425	21	87.5	440	4	US-09-548-376D-53	Sequence 53, Appl	492	21	87.5	476	4	US-09-548-376D-6	Sequence 6, Appl
426	21	87.5	444	4	US-09-724-566A-67	Sequence 67, Appl	493	21	87.5	476	4	US-09-548-376D-73	Sequence 73, Appl
427	21	87.5	444	4	US-09-403-269-13	Sequence 13, Appl	494	21	87.5	476	4	US-09-548-376D-73	Sequence 73, Appl
428	21	87.5	445	2	US-08-679-635A-3	Sequence 3, Appl	495	21	87.5	476	4	US-09-794-927A-6	Sequence 6, Appl
429	21	87.5	445	3	US-09-419-163-3	Sequence 3, Appl	496	21	87.5	476	4	US-09-548-373D-6	Sequence 6, Appl
430	21	87.5	446	3	US-09-199-637A-267	Sequence 267, App	497	21	87.5	476	4	US-09-548-373D-73	Sequence 73, Appl
431	21	87.5	446	4	US-09-548-372D-22	Sequence 22, Appl	498	21	87.5	476	4	US-09-869-414-6	Sequence 6, Appl
432	21	87.5	446	4	US-09-548-367D-22	Sequence 22, Appl	500	21	87.5	476	4	US-09-548-366F-6	Sequence 6, Appl
433	21	87.5	446	4	US-09-551-853D-22	Sequence 22, Appl	501	21	87.5	477	2	US-08-770-544-16	Sequence 16, Appl
434	21	87.5	446	4	US-09-416-901B-22	Sequence 22, Appl	502	21	87.5	477	4	US-09-579-259-16	Sequence 16, Appl
435	21	87.5	446	4	US-09-548-376D-22	Sequence 22, Appl	503	21	87.5	477	4	US-09-650-324A-16	Sequence 16, Appl
436	21	87.5	446	4	US-09-794-927A-22	Sequence 22, Appl	504	21	87.5	480	4	US-09-724-566A-66	Sequence 66, Appl
437	21	87.5	446	4	US-09-548-373D-22	Sequence 22, Appl	505	21	87.5	482	4	US-09-710-279-1314	Sequence 3826, Ap
438	21	87.5	446	4	US-09-795-847B-22	Sequence 22, Appl	506	21	87.5	485	3	US-09-134-001C-3626	Sequence 26680, A
439	21	87.5	446	4	US-09-869-414-22	Sequence 22, Appl	507	21	87.5	485	4	US-09-252-991A-26680	Sequence 2, Appl
440	21	87.5	446	4	US-09-548-366F-22	Sequence 22, Appl	508	21	87.5	488	4	US-09-266-320D-2	Sequence 2, Appl
441	21	87.5	452	4	US-09-724-566A-59	Sequence 59, Appl	509	21	87.5	488	4	US-09-604-608-2	Sequence 2, Appl
442	21	87.5	453	4	US-09-548-372D-30	Sequence 30, Appl	510	21	87.5	492	3	US-09-342-749-2	Sequence 2, Appl
443	21	87.5	453	4	US-09-548-367D-30	Sequence 30, Appl	511	21	87.5	492	4	US-09-691-840-2	Sequence 2, Appl
444	21	87.5	453	4	US-09-551-853D-30	Sequence 30, Appl	512	21	87.5	493	4	US-09-489-039A-11861	Sequence 11861, A
445	21	87.5	453	4	US-09-416-901B-30	Sequence 30, Appl	513	21	87.5	501	3	US-09-009-191-2	Sequence 2, Appl
446	21	87.5	453	4	US-09-548-376D-30	Sequence 30, Appl	514	21	87.5	501	3	US-09-713-158-2	Sequence 2, Appl
447	21	87.5	453	4	US-09-794-927A-30	Sequence 30, Appl	515	21	87.5	501	4	US-09-548-372D-4	Sequence 4, Appl
448	21	87.5	453	4	US-09-548-373D-30	Sequence 30, Appl	516	21	87.5	501	4	US-09-548-372D-8	Sequence 8, Appl
449	21	87.5	453	4	US-09-795-847B-30	Sequence 30, Appl	517	21	87.5	501	4	US-09-548-367D-8	Sequence 8, Appl
450	21	87.5	453	4	US-09-869-414-30	Sequence 30, Appl	518	21	87.5	501	4	US-09-548-367D-8	Sequence 8, Appl
451	21	87.5	456	4	US-09-724-566A-43	Sequence 43, Appl	519	21	87.5	501	4	US-09-551-853D-8	Sequence 8, Appl
452	21	87.5	457	4	US-09-543-681A-6481	Sequence 6481, Ap	520	21	87.5	501	4	US-09-551-853D-8	Sequence 8, Appl
453	21	87.5	457	4	US-09-248-796A-14536	Sequence 14536, A	521	21	87.5	501	4	US-09-724-566A-2	Sequence 2, Appl
454	21	87.5	457	4	US-09-548-372D-24	Sequence 24, Appl	522	21	87.5	501	4	US-09-724-566A-65	Sequence 65, Appl
455	21	87.5	459	4	US-09-548-367D-24	Sequence 24, Appl	523	21	87.5	501	4	US-09-416-901B-4	Sequence 4, Appl
456	21	87.5	459	4	US-09-548-367D-32	Sequence 32, Appl	524	21	87.5	501	4	US-09-416-901B-8	Sequence 8, Appl
457	21	87.5	459	4	US-09-551-853D-24	Sequence 32, Appl	525	21	87.5	501	4	US-09-548-376D-4	Sequence 4, Appl
458	21	87.5	459	4	US-09-551-853D-32	Sequence 32, Appl	526	21	87.5	501	4	US-09-548-376D-8	Sequence 8, Appl
459	21	87.5	459	4	US-09-416-901B-24	Sequence 24, Appl	527	21	87.5	501	4	US-09-794-927A-4	Sequence 4, Appl
460	21	87.5	459	4	US-09-416-901B-32	Sequence 32, Appl	528	21	87.5	501	4	US-09-794-927A-8	Sequence 8, Appl
461	21	87.5	459	4	US-09-416-901B-32	Sequence 32, Appl	529	21	87.5	501	4	US-09-548-373D-4	Sequence 4, Appl
462	21	87.5	459	4	US-09-548-376D-24	Sequence 24, Appl	530	21	87.5	501	4	US-09-548-373D-8	Sequence 8, Appl
463	21	87.5	459	4	US-09-548-376D-32	Sequence 32, Appl	531	21	87.5	501	4	US-09-548-373D-8	Sequence 8, Appl
464	21	87.5	459	4	US-09-794-927A-24	Sequence 24, Appl	532	21	87.5	501	4	US-09-795-847B-8	Sequence 8, Appl
465	21	87.5	459	4			533	21	87.5	501	4	US-09-869-414-8	Sequence 8, Appl
							534	21	87.5	501	4	US-09-869-414-8	Sequence 8, Appl
							535	21	87.5	501	4	US-09-548-366F-4	Sequence 4, Appl
							536	21	87.5	501	4	US-09-548-366F-8	Sequence 8, Appl
							537	21	87.5	502	2	US-08-459-818-19	Sequence 19, Appl
							538	21	87.5	502	2	US-08-889-666-19	Sequence 19, Appl

539	21	87.5	502	2	US-08-465-078-19	Sequence 19, Appl	612	21	87.5	641	4	US-09-328-352-6991	Sequence 6991, Ap
540	21	87.5	502	2	US-08-725-776-19	Sequence 19, Appl	613	21	87.5	641	4	US-09-107-532A-6094	Sequence 6094, Ap
541	21	87.5	502	2	US-08-488-062-19	Sequence 19, Appl	614	21	87.5	641	4	US-09-543-681A-4359	Sequence 4359, Ap
542	21	87.5	503	3	US-09-134-001C-4511	Sequence 4511, Ap	615	21	87.5	641	4	US-09-134-000C-5865	Sequence 5865, Ap
543	21	87.5	503	3	US-09-604-608-3	Sequence 3, Appl	616	21	87.5	642	4	US-09-543-681A-4606	Sequence 4606, Ap
544	21	87.5	504	4	US-09-270-767-43244	Sequence 43244, A	617	21	87.5	647	4	US-09-543-681A-5240	Sequence 5240, Ap
545	21	87.5	516	4	US-09-248-796A-17166	Sequence 17166, A	618	21	87.5	649	4	US-09-543-681A-18996	Sequence 18996, A
546	21	87.5	517	4	US-09-248-796A-14631	Sequence 14631, A	619	21	87.5	653	4	US-09-540-236-2345	Sequence 2345, Ap
547	21	87.5	518	4	US-09-248-796A-18238	Sequence 18238, A	620	21	87.5	689	4	US-08-778-570B-16	Sequence 16, Appl
548	21	87.5	522	1	US-08-639-237-2	Sequence 2, Appl	621	21	87.5	689	4	US-09-059-584-16	Sequence 16, Appl
549	21	87.5	522	1	US-08-975-405-2	Sequence 2, Appl	622	21	87.5	693	3	US-09-564-805-234	Sequence 234, App
550	21	87.5	526	4	US-09-252-991A-29246	Sequence 29246, A	623	21	87.5	698	4	US-09-489-847-200	Sequence 200, App
551	21	87.5	526	4	US-09-328-352-8024	Sequence 8024, Ap	624	21	87.5	700	4	US-09-107-532A-5094	Sequence 5094, Ap
552	21	87.5	527	4	US-09-328-352-7618	Sequence 7618, Ap	625	21	87.5	705	4	US-09-270-767-44461	Sequence 44461, A
553	21	87.5	527	4	US-09-328-352-7733	Sequence 7733, Ap	626	21	87.5	709	4	US-08-778-570B-15	Sequence 15, Appl
554	21	87.5	528	4	US-09-356-806-8	Sequence 8, Appl	627	21	87.5	709	4	US-09-059-584-15	Sequence 15, Appl
555	21	87.5	530	3	US-08-975-762-73	Sequence 73, Appl	628	21	87.5	710	4	US-09-252-991A-24946	Sequence 24946, A
556	21	87.5	530	3	US-09-295-028-73	Sequence 73, Appl	629	21	87.5	712	4	US-09-059-584-51	Sequence 51, Appl
557	21	87.5	530	3	US-09-106-582-73	Sequence 73, Appl	630	21	87.5	717	6	5262177-5	Patent No. 5262177
558	21	87.5	530	3	US-09-159-469-73	Sequence 73, Appl	631	21	87.5	719	1	US-08-082-849B-31	Sequence 31, Appl
559	21	87.5	530	4	US-09-693-542-73	Sequence 73, Appl	632	21	87.5	719	2	US-08-520-933-3	Sequence 3, Appl
560	21	87.5	530	4	US-09-693-542-73	Sequence 73, Appl	633	21	87.5	719	4	US-09-285-040-3	Sequence 3, Appl
561	21	87.5	532	3	US-09-117-853-2	Sequence 2, Appl	634	21	87.5	719	5	PCT-US94-01624-31	Sequence 31, Appl
562	21	87.5	532	3	US-09-911-154-2	Sequence 2, Appl	635	21	87.5	733	4	US-09-248-796A-16565	Sequence 16565, A
563	21	87.5	532	4	US-09-485-529-2	Sequence 2, Appl	636	21	87.5	735	1	US-08-021-601-4	Sequence 4, Appl
564	21	87.5	532	4	US-09-911-514-2	Sequence 2, Appl	637	21	87.5	735	1	US-08-082-849B-4	Sequence 4, Appl
565	21	87.5	540	4	US-09-252-991A-22002	Sequence 22002, A	638	21	87.5	735	5	PCT-US94-01624-4	Sequence 4, Appl
566	21	87.5	546	3	US-09-066-047-2	Sequence 2, Appl	639	21	87.5	736	4	US-09-350-729A-7	Sequence 7, Appl
567	21	87.5	548	3	US-09-352-990-12	Sequence 12, Appl	640	21	87.5	736	6	5262177-2	Patent No. 5262177
568	21	87.5	549	2	US-08-791-887-2	Sequence 2, Appl	641	21	87.5	739	2	US-08-836-943-2	Sequence 2, Appl
569	21	87.5	549	4	US-09-146-084-2	Sequence 2, Appl	642	21	87.5	739	2	US-09-540-236-3594	Sequence 3594, Ap
570	21	87.5	557	4	US-09-10-279-1486	Sequence 1486, Ap	643	21	87.5	739	4	US-09-540-236-12856	Sequence 12856, A
571	21	87.5	558	3	US-09-134-001C-4178	Sequence 4178, Ap	644	21	87.5	747	4	US-09-489-039A-33073	Sequence 33073, A
572	21	87.5	560	4	US-09-489-039A-9311	Sequence 9311, Ap	645	21	87.5	751	4	US-09-252-991A-33073	Sequence 33073, A
573	21	87.5	560	4	US-09-252-991A-27301	Sequence 27301, A	646	21	87.5	763	4	US-09-350-729A-5	Sequence 5, Appl
574	21	87.5	566	4	US-09-252-991A-26648	Sequence 26648, A	647	21	87.5	764	4	US-09-350-729A-6	Sequence 6, Appl
575	21	87.5	567	4	US-09-248-796A-16851	Sequence 16851, A	648	21	87.5	770	4	US-09-252-991A-25267	Sequence 25267, A
576	21	87.5	569	4	US-08-311-731A-106	Sequence 106, App	649	21	87.5	774	3	US-09-009-191-4	Sequence 4, Appl
577	21	87.5	569	4	US-09-350-729A-8	Sequence 8, Appl	650	21	87.5	779	3	US-09-564-805-235	Sequence 235, App
578	21	87.5	574	2	US-08-907-166-12	Sequence 12, Appl	651	21	87.5	788	1	US-08-572-225-1	Sequence 1, Appl
579	21	87.5	574	4	US-09-391-340-12	Sequence 12, Appl	652	21	87.5	806	4	US-09-976-594-321	Sequence 321, App
580	21	87.5	578	4	US-09-248-796A-14802	Sequence 14802, A	653	21	87.5	812	4	US-09-270-767-57052	Sequence 57052, A
581	21	87.5	580	4	US-09-248-796A-16405	Sequence 16405, A	654	21	87.5	812	4	US-09-270-767-57403	Sequence 57403, A
582	21	87.5	583	2	US-08-616-392C-4	Sequence 4, Appl	655	21	87.5	846	1	US-08-149-103-3	Sequence 3, Appl
583	21	87.5	590	3	US-08-975-762-64	Sequence 64, Appl	656	21	87.5	846	1	US-08-451-883-3	Sequence 3, Appl
584	21	87.5	590	3	US-09-295-028-64	Sequence 64, Appl	657	21	87.5	849	4	US-09-248-796A-15566	Sequence 15566, A
585	21	87.5	590	3	US-09-106-582-64	Sequence 64, Appl	658	21	87.5	859	4	US-09-107-532A-6829	Sequence 6829, Ap
586	21	87.5	590	4	US-09-159-469-64	Sequence 64, Appl	659	21	87.5	869	4	US-09-543-681A-4814	Sequence 4814, Ap
587	21	87.5	590	4	US-09-693-542-64	Sequence 64, Appl	660	21	87.5	873	1	US-08-393-734-2	Sequence 2, Appl
588	21	87.5	593	4	US-09-328-352-4462	Sequence 4462, Ap	661	21	87.5	873	3	US-08-894-489-2	Sequence 2, Appl
589	21	87.5	593	4	US-09-352-990-10	Sequence 10, Appl	662	21	87.5	873	4	US-09-543-681A-7315	Sequence 7315, Ap
590	21	87.5	603	4	US-09-248-796A-17178	Sequence 17178, A	663	21	87.5	900	2	US-08-813-940-4	Sequence 4, Appl
591	21	87.5	604	4	US-09-508-213-3	Sequence 3, Appl	664	21	87.5	903	1	US-08-021-601-12	Sequence 12, Appl
592	21	87.5	607	3	US-08-781-891-75	Sequence 75, Appl	665	21	87.5	903	1	US-08-082-849B-12	Sequence 12, Appl
593	21	87.5	607	4	US-09-328-352-6716	Sequence 6716, Ap	666	21	87.5	903	5	PCT-US94-01624-12	Sequence 12, Appl
594	21	87.5	607	4	US-09-618-166-75	Sequence 75, Appl	667	21	87.5	905	1	US-08-072-574-2	Sequence 2, Appl
595	21	87.5	607	4	US-09-344-882-16	Sequence 16, Appl	668	21	87.5	906	1	US-08-486-270-2	Sequence 2, Appl
596	21	87.5	607	4	US-09-344-882-18	Sequence 18, Appl	669	21	87.5	906	3	US-08-367-264-2	Sequence 2, Appl
597	21	87.5	613	4	US-09-331-568A-25	Sequence 25, Appl	670	21	87.5	906	4	US-09-153-757-2	Sequence 2, Appl
598	21	87.5	619	4	US-09-489-039A-8075	Sequence 8075, Ap	671	21	87.5	906	4	US-09-459-715-2	Sequence 2, Appl
599	21	87.5	621	4	US-09-331-568A-2	Sequence 2, Appl	672	21	87.5	906	5	PCT-US91-09422-17	Sequence 17, Appl
600	21	87.5	621	4	US-09-331-568A-26	Sequence 26, Appl	673	21	87.5	922	4	US-09-198-452A-15	Sequence 15, Appl
601	21	87.5	623	3	US-09-104-068-4	Sequence 4, Appl	674	21	87.5	932	4	US-08-887-534A-45	Sequence 45, Appl
602	21	87.5	623	4	US-09-485-529-7	Sequence 7, Appl	675	21	87.5	932	4	US-09-527-431-45	Sequence 45, Appl
603	21	87.5	625	4	US-09-710-279-2464	Sequence 2464, Ap	676	21	87.5	932	4	US-09-446-861-45	Sequence 45, Appl
604	21	87.5	626	4	US-09-331-568A-27	Sequence 27, Appl	677	21	87.5	943	2	US-08-808-982-7	Sequence 7, Appl
605	21	87.5	630	4	US-09-485-529-8	Sequence 8, Appl	678	21	87.5	943	3	US-09-306-902A-7	Sequence 7, Appl
606	21	87.5	631	3	US-09-134-001C-4605	Sequence 4605, Ap	679	21	87.5	945	4	US-10-140-002-146	Sequence 146, App
607	21	87.5	635	4	US-09-583-110-4021	Sequence 4021, Ap	680	21	87.5	979	3	US-08-514-213A-2	Sequence 2, Appl
608	21	87.5	635	4	US-09-543-681A-7499	Sequence 7499, Ap	681	21	87.5	979	4	US-09-015-399-5	Sequence 5, Appl
609	21	87.5	637	3	US-09-104-068-2	Sequence 2, Appl	682	21	87.5	985	4	US-09-248-796A-16090	Sequence 16090, A
610	21	87.5	637	4	US-09-583-110-4865	Sequence 4865, Ap	683	21	87.5	986	3	US-08-742-757-4	Sequence 4, Appl
611	21	87.5	640	4	US-09-724-623-96	Sequence 96, Appl	684	21	87.5	986	4	US-09-850-048A-4	Sequence 4, Appl



831	20	83.3	157	4	US-09-134-000C-3711	Sequence 3711, Ap	904	20	83.3	311	4	US-09-393-858-20	Sequence 20, Appl
832	20	83.3	163	3	US-09-134-001C-4131	Sequence 4131, Ap	905	20	83.3	311	4	US-09-248-796A-16692	Sequence 16692, A
833	20	83.3	163	4	US-09-328-352-5655	Sequence 5655, Ap	906	20	83.3	311	4	US-10-190-279-20	Sequence 20, Appl
834	20	83.3	163	4	US-09-761-782-2	Sequence 2, Appli	907	20	83.3	314	4	US-09-248-796A-18201	Sequence 18201, A
835	20	83.3	169	3	US-09-134-001C-5390	Sequence 5390, Ap	908	20	83.3	315	4	US-09-248-796A-19221	Sequence 19221, A
836	20	83.3	170	4	US-09-732-210-1074	Sequence 1074, Ap	909	20	83.3	321	2	US-08-780-370A-2	Sequence 2, Appli
837	20	83.3	174	4	US-09-107-532A-5979	Sequence 5979, Ap	910	20	83.3	321	4	US-09-251-330-2	Sequence 2, Appli
838	20	83.3	174	4	US-09-489-039A-12363	Sequence 12363, A	911	20	83.3	323	4	US-09-248-796A-15432	Sequence 15432, A
839	20	83.3	178	4	US-09-134-000C-3409	Sequence 3409, Ap	912	20	83.3	324	4	US-09-252-991A-28877	Sequence 28877, A
840	20	83.3	178	4	US-09-540-236-2174	Sequence 2174, Ap	913	20	83.3	325	4	US-09-489-039A-12379	Sequence 12379, A
841	20	83.3	178	4	US-09-248-796A-27780	Sequence 27780, A	914	20	83.3	327	4	US-09-489-039A-14005	Sequence 14005, A
842	20	83.3	186	4	US-09-325-932A-72	Sequence 72, Appl	915	20	83.3	329	4	US-09-248-796A-16725	Sequence 16725, A
843	20	83.3	200	4	US-09-252-991A-33039	Sequence 33039, A	916	20	83.3	330	4	US-09-252-991A-23378	Sequence 23378, A
844	20	83.3	201	4	US-09-489-039A-7394	Sequence 7394, Ap	917	20	83.3	332	4	US-09-489-039A-12580	Sequence 12580, A
845	20	83.3	203	4	US-09-134-000C-6202	Sequence 6202, Ap	918	20	83.3	334	4	US-09-489-039A-9368	Sequence 9368, Ap
846	20	83.3	205	3	US-09-134-001C-4766	Sequence 4766, Ap	919	20	83.3	346	4	US-09-252-991A-26418	Sequence 26418, A
847	20	83.3	207	4	US-09-540-236-2475	Sequence 2475, Ap	920	20	83.3	347	3	US-09-134-001C-5508	Sequence 5508, Ap
848	20	83.3	208	3	US-09-094-103-4	Sequence 4, Appli	921	20	83.3	348	3	US-09-134-001C-4742	Sequence 4742, Ap
849	20	83.3	208	4	US-09-328-352-5212	Sequence 5212, Ap	922	20	83.3	351	2	US-08-843-530B-31	Sequence 31, Appl
850	20	83.3	209	4	US-09-710-279-688	Sequence 688, App	923	20	83.3	351	4	US-09-636-728-28	Sequence 28, Appl
851	20	83.3	209	4	US-09-710-279-2236	Sequence 2236, Ap	924	20	83.3	352	4	US-09-252-991A-25209	Sequence 25209, A
852	20	83.3	216	3	US-08-824-692-24	Sequence 24, Appl	925	20	83.3	354	2	US-08-843-530B-30	Sequence 30, Appl
853	20	83.3	216	3	US-08-914-375C-17	Sequence 17, Appl	926	20	83.3	354	4	US-09-636-728-27	Sequence 27, Appl
854	20	83.3	216	3	US-08-914-375C-19	Sequence 19, Appl	927	20	83.3	354	4	US-08-586-165-7	Sequence 7, Appli
855	20	83.3	216	3	US-08-914-375C-20	Sequence 20, Appl	928	20	83.3	363	3	US-08-586-165-7	Sequence 7, Appli
856	20	83.3	222	1	US-08-126-593A-2	Sequence 2, Appli	929	20	83.3	364	4	US-09-252-991A-31716	Sequence 31716, A
857	20	83.3	222	1	US-08-454-039A-2	Sequence 2, Appli	930	20	83.3	366	1	US-08-700-359-22	Sequence 22, Appl
858	20	83.3	225	4	US-09-328-352-6504	Sequence 6504, Ap	931	20	83.3	367	4	US-09-543-681A-4184	Sequence 4184, Ap
859	20	83.3	229	4	US-09-489-039A-8347	Sequence 8347, Ap	932	20	83.3	371	3	US-08-586-165-3	Sequence 3, Appli
860	20	83.3	231	2	US-08-286-819A-12	Sequence 12, Appl	933	20	83.3	372	3	US-08-586-165-5	Sequence 5, Appli
861	20	83.3	231	2	US-08-286-819A-43	Sequence 43, Appl	934	20	83.3	372	4	US-09-489-039A-12516	Sequence 12516, A
862	20	83.3	231	3	US-08-980-357-12	Sequence 12, Appl	935	20	83.3	385	4	US-09-107-532A-6173	Sequence 6173, Ap
863	20	83.3	231	3	US-08-980-357-43	Sequence 43, Appl	936	20	83.3	385	4	US-09-248-796A-23097	Sequence 23097, A
864	20	83.3	232	3	US-09-094-103-2	Sequence 2, Appli	937	20	83.3	387	4	US-09-563-269-2	Sequence 2, Appli
865	20	83.3	232	3	US-09-583-110-5010	Sequence 5010, Ap	938	20	83.3	387	4	US-09-643-596B-136	Sequence 136, App
866	20	83.3	239	3	US-08-462-436-2	Sequence 2, Appli	939	20	83.3	389	4	US-09-248-796A-17335	Sequence 17335, A
867	20	83.3	239	3	US-08-463-275-2	Sequence 2, Appli	940	20	83.3	390	4	US-09-107-532A-6615	Sequence 6615, Ap
868	20	83.3	239	4	US-08-640-877-2	Sequence 2, Appli	941	20	83.3	392	3	US-08-861-774E-92	Sequence 92, Appl
869	20	83.3	239	4	US-09-799-576A-2	Sequence 2, Appli	942	20	83.3	396	3	US-08-861-774E-84	Sequence 84, Appl
870	20	83.3	239	4	US-09-799-540-2	Sequence 2, Appli	943	20	83.3	400	4	US-09-543-681A-4952	Sequence 4952, Ap
871	20	83.3	240	3	US-08-824-692-23	Sequence 23, Appl	944	20	83.3	401	4	US-09-489-039A-11969	Sequence 11969, A
872	20	83.3	240	4	US-09-583-110-3479	Sequence 3479, Ap	945	20	83.3	404	4	US-09-543-681A-5169	Sequence 5169, Ap
873	20	83.3	242	4	US-09-270-767-33153	Sequence 33153, A	946	20	83.3	404	4	US-09-489-039A-12039	Sequence 12039, A
874	20	83.3	242	4	US-09-270-767-48370	Sequence 48370, A	947	20	83.3	405	4	US-09-540-236-3547	Sequence 3547, Ap
875	20	83.3	244	4	US-09-252-991A-29762	Sequence 29762, A	948	20	83.3	406	4	US-08-311-731A-12	Sequence 12, Appl
876	20	83.3	244	4	US-09-543-681A-7039	Sequence 7039, Ap	949	20	83.3	406	4	US-09-489-039A-10003	Sequence 10003, A
877	20	83.3	258	4	US-09-248-796A-21208	Sequence 21208, A	950	20	83.3	409	2	US-08-743-130A-2	Sequence 2, Appli
878	20	83.3	261	4	US-10-101-464A-691	Sequence 691, App	951	20	83.3	409	2	US-08-743-130A-39	Sequence 39, Appl
879	20	83.3	269	4	US-09-583-110-4586	Sequence 4586, Ap	952	20	83.3	426	4	US-09-602-472A-59	Sequence 59, Appl
880	20	83.3	269	4	US-09-710-279-1158	Sequence 1158, Ap	953	20	83.3	428	4	US-09-248-796A-18691	Sequence 18691, A
881	20	83.3	270	4	US-09-248-796A-18801	Sequence 18801, A	954	20	83.3	430	4	US-09-328-352-5931	Sequence 5931, Ap
882	20	83.3	272	4	US-09-710-279-1812	Sequence 1812, Ap	955	20	83.3	432	1	US-08-317-880-2	Sequence 2, Appli
883	20	83.3	279	3	US-09-134-001C-4878	Sequence 4878, Ap	956	20	83.3	452	2	US-08-782-396-2	Sequence 2, Appli
884	20	83.3	279	4	US-09-252-991A-31984	Sequence 31984, A	957	20	83.3	452	4	US-09-540-236-370	Sequence 370, Ap
885	20	83.3	280	4	US-09-489-039A-13968	Sequence 13968, A	958	20	83.3	455	4	US-08-583-110-3993	Sequence 3993, Ap
886	20	83.3	281	4	US-08-861-451A-22	Sequence 22, Appl	959	20	83.3	458	4	US-09-583-110-3651	Sequence 3651, Ap
887	20	83.3	284	4	US-09-252-991A-26484	Sequence 26484, A	960	20	83.3	464	4	US-09-351-150A-9	Sequence 9, Appli
888	20	83.3	286	4	US-09-543-681A-5287	Sequence 5287, Ap	961	20	83.3	470	3	US-09-004-838-44	Sequence 44, Appl
889	20	83.3	286	4	US-09-248-796A-17854	Sequence 17854, A	962	20	83.3	472	3	US-09-004-838-43	Sequence 43, Appl
890	20	83.3	290	4	US-09-643-596B-144	Sequence 144, App	963	20	83.3	473	4	US-09-328-352-7514	Sequence 7514, Ap
891	20	83.3	295	3	US-09-134-001C-4242	Sequence 4242, Ap	964	20	83.3	474	4	US-09-489-039A-14147	Sequence 14147, A
892	20	83.3	295	4	US-09-538-092-607	Sequence 607, App	965	20	83.3	475	4	US-09-248-796A-17355	Sequence 17355, A
893	20	83.3	298	4	US-09-248-796A-18159	Sequence 18159, A	966	20	83.3	481	4	US-09-198-452A-1012	Sequence 1012, Ap
894	20	83.3	303	4	US-09-543-681A-7924	Sequence 7924, Ap	967	20	83.3	487	4	US-09-543-681A-7031	Sequence 7031, Ap
895	20	83.3	303	4	US-09-543-681A-7946	Sequence 7946, Ap	968	20	83.3	491	3	US-09-426-568A-2	Sequence 2, Appli
896	20	83.3	304	4	US-09-252-991A-18757	Sequence 18757, A	969	20	83.3	495	4	US-09-538-092-157	Sequence 157, App
897	20	83.3	307	3	US-09-134-001C-3238	Sequence 3238, Ap	970	20	83.3	495	4	US-09-252-991A-24541	Sequence 24541, A
898	20	83.3	307	4	US-09-583-110-3427	Sequence 3427, Ap	971	20	83.3	500	3	US-09-134-001C-5280	Sequence 5280, Ap
899	20	83.3	310	3	US-09-372-934-4	Sequence 4, Appli	972	20	83.3	501	3	US-09-134-001C-3978	Sequence 3978, Ap
900	20	83.3	310	4	US-09-766-898-4	Sequence 4, Appli	973	20	83.3	507	4	US-09-198-452A-174	Sequence 174, App
901	20	83.3	310	4	US-09-766-898-4	Sequence 4, Appli	974	20	83.3	514	4	US-09-134-000C-5471	Sequence 5471, Ap
902	20	83.3	310	4	US-09-637-040C-4	Sequence 4, Appli	975	20	83.3	514	4	US-09-710-279-2682	Sequence 2682, Ap
903	20	83.3	311	3	US-09-564-805-230	Sequence 230, App	976	20	83.3	516	3	US-09-215-694-16	Sequence 16, Appl
							977	20	83.3	516	4	US-09-583-110-4311	Sequence 4311, Ap

977 20 83.3 518 3 US-09-134-001C-4451  
978 20 83.3 521 2 US-08-682-847-4  
979 20 83.3 522 4 US-09-489-039A-10041  
980 20 83.3 525 4 US-09-328-352-4492  
981 20 83.3 527 4 US-09-107-532A-6562  
982 20 83.3 529 3 US-09-433-248A-6  
983 20 83.3 543 4 US-09-595-386-6  
984 20 83.3 543 4 US-09-933-525-6  
985 20 83.3 551 4 US-09-248-796A-17366  
986 20 83.3 566 4 US-09-543-681A-7802  
987 20 83.3 573 4 US-09-252-991A-20056  
988 20 83.3 575 1 US-08-348-920-1  
989 20 83.3 575 1 US-08-348-920-2  
990 20 83.3 575 3 US-09-134-001C-3906  
991 20 83.3 576 4 US-09-328-352-6971  
992 20 83.3 585 4 US-09-489-039A-10279  
993 20 83.3 589 4 US-09-328-352-7504  
994 20 83.3 600 2 US-08-370-156-4  
995 20 83.3 600 3 US-08-814-095-4  
996 20 83.3 600 3 US-08-975-084-1  
997 20 83.3 600 3 US-09-134-001C-3636  
998 20 83.3 602 2 US-08-419-652-6  
999 20 83.3 602 3 US-08-446-100-1  
1000 20 83.3 602 3 US-08-446-100-2

ALIGNMENTS

RESULT 1  
US-08-934-915-30  
; Sequence 30, Application US/08934915  
; Patent No. 5932412  
; GENERAL INFORMATION:  
; APPLICANT: DILLNER, JOAKIM  
; APPLICANT: DILLNER, LENA  
; APPLICANT: CHENG, HWEE-MING  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
; TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
; NUMBER OF SEQUENCES: 193  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MASON & ASSOCIATES, P.A.  
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
; CITY: CLEARWATER  
; STATE: FLORIDA  
; COUNTRY: U.S.A.  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 3.0  
; SOFTWARE: Microsoft Word 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/934,915  
; FILING DATE: 22-SEP-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/949,836  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LOUISE A. Foutch  
; REGISTRATION NUMBER: 37,133  
; REFERENCE/DOCKET NUMBER: 1946.6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 813-538-3800  
; TELEFAX: 813-538-3820  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-934-915-30  
Query Match 91.7%; Score 22; DB 2; Length 20;  
Best Local Similarity 80.0%; Pred. No. 44;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
DB 9 GYTV 13  
RESULT 2  
US-08-934-915-135  
; Sequence 135, Application US/08934915  
; Patent No. 5932412  
; GENERAL INFORMATION:  
; APPLICANT: DILLNER, JOAKIM  
; APPLICANT: DILLNER, LENA  
; APPLICANT: CHENG, HWEE-MING  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
; TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
; NUMBER OF SEQUENCES: 193  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MASON & ASSOCIATES, P.A.  
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
; CITY: CLEARWATER  
; STATE: FLORIDA  
; COUNTRY: U.S.A.  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 3.0  
; SOFTWARE: Microsoft Word 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/934,915  
; FILING DATE: 22-SEP-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/949,836  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LOUISE A. Foutch  
; REGISTRATION NUMBER: 37,133  
; REFERENCE/DOCKET NUMBER: 1946.6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 813-538-3800  
; TELEFAX: 813-538-3820  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 135:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-934-915-135  
Query Match 91.7%; Score 22; DB 2; Length 20;  
Best Local Similarity 80.0%; Pred. No. 44;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVE 5  
DB 9 GYTV 13  
RESULT 3  
US-09-621-976-7003  
; Sequence 7003, Application US/09621976



; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7003  
; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-7003

Query Match 91.7%; Score 22; DB 4; Length 54;  
Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 45 GYTV 49

RESULT 4  
US-09-270-767-33535  
; Sequence 33535, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 33535  
; LENGTH: 60  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-33535

Query Match 91.7%; Score 22; DB 4; Length 60;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 29 GYSVE 33

RESULT 5  
US-09-270-767-48752  
; Sequence 48752, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48752  
; LENGTH: 60  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-48752

Query Match 91.7%; Score 22; DB 4; Length 60;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GYXVE 5  
|||  
Db 29 GYSVE 33

RESULT 6  
US-08-454-557C-36  
; Sequence 36, Application US/08454557C  
; Patent No. 5830670  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; APPLICANT: Wands, Jack R.  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; TITLE OF INVENTION: of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,557C  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ludwig, Steven R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0609.3840003  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 68 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-454-557C-36

Query Match 91.7%; Score 22; DB 2; Length 68;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 43 GYAVE 47

RESULT 7  
US-08-340-426D-36  
; Sequence 36, Application US/08340426D  
; Patent No. 5948634  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; APPLICANT: Wands, Jack R.  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; TITLE OF INVENTION: of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/340,426D  
FILING DATE: 14-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, Steven R.  
REGISTRATION NUMBER: 36,203  
REFERENCE/DOCKET NUMBER: 0609.38400002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-340-426D-36

Query Match 91.7%; Score 22; DB 2; Length 68;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
DB 43 GYAVE 47

RESULT 8  
US-08-450-673C-36  
Sequence 36, Application US/08450673C  
Patent No. 5948888  
GENERAL INFORMATION:  
APPLICANT: de la Monte, Suzanne  
APPLICANT: Wands, Jack R.  
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
TITLE OF INVENTION: of Alzheimer's Disease  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,673C  
FILING DATE: 30-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, Steven R.  
REGISTRATION NUMBER: 36,203  
REFERENCE/DOCKET NUMBER: 0609.38400004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-450-673C-36

Query Match 91.7%; Score 22; DB 2; Length 68;

Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
DB 43 GYAVE 47

RESULT 9  
PCT-US95-17111A-36  
Sequence 36, Application PC/TUS9517111A  
GENERAL INFORMATION:  
APPLICANT: de la Monte, Suzanne  
APPLICANT: Wands, Jack R.  
TITLE OF INVENTION: Neural Thread Protein Gene Expression and  
TITLE OF INVENTION: Detection of Alzheimer's Disease  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/17111A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/340,426  
FILING DATE: 14-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, Steven R.  
REGISTRATION NUMBER: 36,203  
REFERENCE/DOCKET NUMBER: 0609.38400002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US95-17111A-36

Query Match 91.7%; Score 22; DB 5; Length 68;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
DB 43 GYAVE 47

RESULT 10  
US-08-447-173A-58  
Sequence 58, Application US/08447173A  
Patent No. 5843730  
GENERAL INFORMATION:  
APPLICANT: Wain-Hobson, Simon  
APPLICANT: Pezo, Valerie  
TITLE OF INVENTION: Method For Hypermutagenesis Of  
TITLE OF INVENTION: Nucleotides  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I. Street, N.W., Suite 700

;  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,173A  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03495.0142-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)408-4000  
; TELEFAX: (202)408-4400  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 78 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-447-173A-58

Query Match 91.7%; Score 22; DB 2; Length 78;  
Best Local Similarity 80.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 54 GYAVE 58

RESULT 11  
US-08-447-173A-61  
; Sequence 61, Application US/08447173A  
; Patent No. 5843730  
; GENERAL INFORMATION:  
; APPLICANT: Wain-Hobson, Simon  
; APPLICANT: Pezo, Valerie  
; TITLE OF INVENTION: Method For Hypermutagenesis Of  
; TITLE OF INVENTION: Nucleotides  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESS: Dunner  
; STREET: 1300 I. Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,173A  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03495.0142-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)408-4000  
; TELEFAX: (202)408-4400  
; INFORMATION FOR SEQ ID NO: 61:

;  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 78 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-447-173A-61

Query Match 91.7%; Score 22; DB 2; Length 78;  
Best Local Similarity 80.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 54 GYAVE 58

RESULT 12  
US-08-241-853-18  
; Sequence 18, Application US/08241853  
; Patent No. 5693488  
; GENERAL INFORMATION:  
; APPLICANT: Fang, Kathy S.  
; APPLICANT: Hanafusa, Hideaburo  
; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE  
; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/241,853  
; FILING DATE: 12-MAY-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-078  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 93 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; US-08-241-853-18

Query Match 91.7%; Score 22; DB 1; Length 93;  
Best Local Similarity 80.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 34 GYTVE 38

RESULT 13  
US-08-850-917-18

```
; Sequence 18, Application US/08850917
; Patent No. 5854045
; GENERAL INFORMATION:
; APPLICANT: Fang, Kathy S.
; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,917
; FILING DATE: 02-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/241,853
; FILING DATE: 12-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-850-917-18

Query Match          91.7%; Score 22; DB 2; Length 93;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      34 GYTV 38

RESULT 14
US-09-609-324A-8
; Sequence 8, Application US/09609324A
; Patent No. RE37582
; GENERAL INFORMATION:
; APPLICANT: CERRETTI, Douglas P.
; TITLE OF INVENTION: CYTOKINE DESIGNATED LERK-6
; FILE REFERENCE: A7772
; CURRENT APPLICATION NUMBER: US/09/609,324A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 08/920,440
; PRIOR FILING DATE: 1997-08-29
; PRIOR APPLICATION NUMBER: 08/538,709
; PRIOR FILING DATE: 1995-10-03
; PRIOR APPLICATION NUMBER: 08/318,393
; PRIOR FILING DATE: 1994-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 8
; LENGTH: 104
; TYPE: PRT
; ORGANISM: LERK-6
; US-09-609-324A-8

Query Match          91.7%; Score 22; DB 1; Length 104;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      11 GYTV 15

RESULT 15
US-08-920-440B-8
; Sequence 8, Application US/08920440B
; Patent No. 5919905
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,440B
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-920-440B-8

Query Match          91.7%; Score 22; DB 2; Length 104;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      11 GYTV 15

RESULT 16
US-09-173-492-8
; Sequence 8, Application US/09173492
; Patent No. 6194172
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
```

STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: System 7.6  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/173,492  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/920,440  
FILING DATE: 29-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Henry, Janis C.  
REGISTRATION NUMBER: 34,347  
REFERENCE/DOCKET NUMBER: 2826-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 104 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-173-492-8

Query Match 91.7%; Score 22; DB 3; Length 104;  
Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 11 GYTVE 15

RESULT 17  
US-09-173-133-8  
; Sequence 8, Application US/09173133  
; Patent No. 6232447  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated LERK-6  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: System 7.6  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/173,133  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,440  
; FILING DATE: 29-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2826-B  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 104 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-173-133-8

Query Match 91.7%; Score 22; DB 3; Length 104;  
Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 11 GYTVE 15

RESULT 18  
US-09-165-533-8  
; Sequence 8, Application US/09165533  
; Patent No. 6268482  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated as LERK-6  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/165,533  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/538,709  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Malaska, Stephen L.  
; REGISTRATION NUMBER: 32,655  
; REFERENCE/DOCKET NUMBER: 2826  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 104 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-165-533-8

Query Match 91.7%; Score 22; DB 3; Length 104;  
Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 11 GYTVE 15

RESULT 19  
US-09-580-236A-8  
; Sequence 8, Application US/09580236A

```

; Patent No. 6472174
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/580,236A
; FILING DATE: 26-May-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-580-236A-8

Query Match          91.7%; Score 22; DB 4; Length 104;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      11 GYTV 15

RESULT 20
US-09-538-092-368
; Sequence 368, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 368
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YGR269W
US-09-538-092-368

Query Match          91.7%; Score 22; DB 4; Length 104;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      11 GYTV 15

; Patent No. 6472174
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/580,236A
; FILING DATE: 26-May-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-580-236A-8

Query Match          91.7%; Score 22; DB 4; Length 104;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      11 GYTV 15

RESULT 20
US-09-538-092-368
; Sequence 368, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 368
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YGR269W
US-09-538-092-368

Query Match          91.7%; Score 22; DB 4; Length 108;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      44 GYTV 48

RESULT 21
US-09-142-469-5
; Sequence 5, Application US/09142469
; Patent No. 6140094
; GENERAL INFORMATION:
; APPLICANT: LOFFLER, Fridolin
; APPLICANT: JUNGSCHAFFER, Gerald
; APPLICANT: KHANH, Quoc Nguyen
; APPLICANT: SCHUSTER, Erwin
; APPLICANT: SPROSSLER, Bruno
; APPLICANT: WOLF, Sabine
; TITLE OF INVENTION: PROTEIN WITH PHOSPHOLIPASE ACTIVITY
; FILE REFERENCE: 015200-056
; CURRENT APPLICATION NUMBER: US/09/142,469
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/EP98/00081
; EARLIER FILING DATE: 1998-01-08
; EARLIER APPLICATION NUMBER: DE 197 01 348.1
; EARLIER FILING DATE: 1997-01-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Aspergillus lysophospholipase
US-09-142-469-5

Query Match          91.7%; Score 22; DB 3; Length 112;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      107 GYSVE 111

RESULT 22
US-09-134-000C-4812
; Sequence 4812, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4812
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4812

Query Match          91.7%; Score 22; DB 4; Length 113;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      6 GYSVE 10
```

```
RESULT 23
US-09-134-000C-4154
; Sequence 4154, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4154
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4154

Query Match          91.7%; Score 22; DB 4; Length 140;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 111 GYTV 115

RESULT 24
US-09-252-991A-32985
; Sequence 32985, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32985
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32985

Query Match          91.7%; Score 22; DB 4; Length 143;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 10 GYAVE 14

RESULT 25
US-09-252-991A-25087
; Sequence 25087, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
```

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25087
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25087

Query Match          91.7%; Score 22; DB 4; Length 161;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 54 GYSVE 58

RESULT 26
US-09-134-000C-4743
; Sequence 4743, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4743
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4743

Query Match          91.7%; Score 22; DB 4; Length 164;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 64 GYAVE 68

RESULT 27
US-09-609-324A-2
; Sequence 2, Application US/09609324A
; Patent No. RE37582
; GENERAL INFORMATION:
; APPLICANT: CERRETTI, Douglas P.
; TITLE OF INVENTION: CYTOKINE DESIGNATED LERK-6
; FILE REFERENCE: A7772
; CURRENT APPLICATION NUMBER: US/09/609,324A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 08/920,440
; PRIOR FILING DATE: 1997-08-29
; PRIOR APPLICATION NUMBER: 08/538,709
; PRIOR FILING DATE: 1995-10-03
; PRIOR APPLICATION NUMBER: 08/318,393
; PRIOR FILING DATE: 1994-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 184
; TYPE: PRT
; ORGANISM: LERK-6
```

US-09-609-324A-2

Query Match 91.7%; Score 22; DB 1; Length 184;  
Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 29 GYTVE 33

RESULT 28

US-08-920-440B-2  
; Sequence 2, Application US/08920440B

; Patent No. 5919905

; GENERAL INFORMATION:

; APPLICANT: Cerretti, Douglas P.

; TITLE OF INVENTION: Cytokine Designated LERK-6

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Immunex Corporation

; STREET: 51 University Street

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Apple Power Macintosh

; OPERATING SYSTEM: System 7.6

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/920,440B

; FILING DATE: 29-AUG-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Henry, Janis C.

; REGISTRATION NUMBER: 34,347

; REFERENCE/DOCKET NUMBER: 2826-B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 587-0430

; TELEFAX: (206) 233-0644

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 184 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-920-440B-2

Query Match 91.7%; Score 22; DB 2; Length 184;

Best Local Similarity 80.0%; Pred. No. 4.4e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 29 GYTVE 33

RESULT 29

US-09-173-492-2

; Sequence 2, Application US/09173492

; Patent No. 6194172

; GENERAL INFORMATION:

; APPLICANT: Cerretti, Douglas P.

; TITLE OF INVENTION: Cytokine Designated LERK-6

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Immunex Corporation

; STREET: 51 University Street

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: System 7.6  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/173,492  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,440  
; FILING DATE: 29-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2826-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 184 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-173-492-2

Query Match 91.7%; Score 22; DB 3; Length 184;

Best Local Similarity 80.0%; Pred. No. 4.4e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 29 GYTVE 33

RESULT 30

US-09-173-133-2

; Sequence 2, Application US/09173133

; Patent No. 6232447

; GENERAL INFORMATION:

; APPLICANT: Cerretti, Douglas P.

; TITLE OF INVENTION: Cytokine Designated LERK-6

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Immunex Corporation

; STREET: 51 University Street

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Apple Power Macintosh

; OPERATING SYSTEM: System 7.6

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/173,133

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/920,440

; FILING DATE: 29-AUG-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Henry, Janis C.

; REGISTRATION NUMBER: 34,347

; REFERENCE/DOCKET NUMBER: 2826-B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 587-0430

; TELEFAX: (206) 233-0644

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:



```
;
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-173-133-2
Query Match 91.7%; Score 22; DB 3; Length 184;
Best Local Similarity 80.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 29 GYTVE 33

RESULT 31
US-09-165-533-2
; Sequence 2, Application US/09165533
; Patent No. 6268482
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated as LERK-6
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/165,533
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/538,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Malaska, Stephen L.
; REGISTRATION NUMBER: 32,655
; REFERENCE/DOCKET NUMBER: 2826
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-165-533-2
Query Match 91.7%; Score 22; DB 3; Length 184;
Best Local Similarity 80.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 29 GYTVE 33

RESULT 32
US-09-580-236A-2
; Sequence 2, Application US/09580236A
; Patent No. 6472174
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
```

```
;
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/580,236A
; FILING DATE: 26-May-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-580-236A-2
Query Match 91.7%; Score 22; DB 4; Length 184;
Best Local Similarity 80.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 29 GYTVE 33

RESULT 33
PCT-US95-12779-2
; Sequence 2, Application PC/TUS9512779
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/12779
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Malaska, Stephen L.
; REGISTRATION NUMBER: 32,655
; REFERENCE/DOCKET NUMBER: 2826-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
```

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 184 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-12779-2

Query Match 91.7%; Score 22; DB 5; Length 184;  
Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 29 GYTV 33

RESULT 34  
PCT-US95-15781-2

; Sequence 2, Application PC/TUS9515781  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated Lerk-7  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/15781  
; FILING DATE: 05-DEC-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/351,025  
; FILING DATE: 06-DEC-1994  
; CLASSIFICATION:  
; APPLICATION NUMBER: US 08/396,946  
; FILING DATE: 01-MAR-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Anderson, Kathryn A.  
; REGISTRATION NUMBER: 32,172  
; REFERENCE/DOCKET NUMBER: 2829-WO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 184 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-15781-2

Query Match 91.7%; Score 22; DB 5; Length 184;  
Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 29 GYTV 33

RESULT 35  
US-09-252-991A-18194  
; Sequence 18194, Application US/09252991A

; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18194  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18194

Query Match 91.7%; Score 22; DB 4; Length 203;  
Best Local Similarity 80.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 171 GYSVE 175

RESULT 36  
US-09-270-767-32818  
; Sequence 32818, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 32818  
; LENGTH: 204  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-32818

Query Match 91.7%; Score 22; DB 4; Length 204;  
Best Local Similarity 80.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 73 GYSVE 77

RESULT 37  
US-09-270-767-48035  
; Sequence 48035, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48035  
; LENGTH: 204  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster



```
; Sequence 6, Application US/09214631
; Patent No. 6413730
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha
; APPLICANT: Mbamalu, Geraldine
; APPLICANT: Pawson, Tony
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR
; TITLE OF INVENTION: TYROSINE KINASES
; FILE REFERENCE: 11757.23USWO
; CURRENT APPLICATION NUMBER: US/09/214,631
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: PCT/CA97/00473
; EARLIER FILING DATE: 1997-07-04
; EARLIER APPLICATION NUMBER: 60/021,272
; EARLIER FILING DATE: 1996-07-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-214-631-6

Query Match          91.7%; Score 22; DB 4; Length 209;
Best Local Similarity 80.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      54 GYTVE 58

RESULT 42
US-08-393-462-2
; Sequence 2, Application US/08393462
; Patent No. 677536
; GENERAL INFORMATION:
; APPLICANT: Flanagan, John G.
; APPLICANT: Cheng, Hwai-Jong
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (Text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,462
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011CPC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-393-462-2
```

```
Query Match          91.7%; Score 22; DB 4; Length 209;
Best Local Similarity 80.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      54 GYTVE 58

RESULT 43
PCT-US95-11869-2
; Sequence 2, Application PC/TUS9511869
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11869
; FILING DATE: 19-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011CPC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-11869-2

Query Match          91.7%; Score 22; DB 5; Length 209;
Best Local Similarity 80.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      54 GYTVE 58

RESULT 44
US-09-609-324A-10
; Sequence 10, Application US/09609324A
; Patent No. RE37582
; GENERAL INFORMATION:
; APPLICANT: CERRETTI, Douglas P.
; TITLE OF INVENTION: CYTOKINE DESIGNATED LERK-6
; FILE REFERENCE: A7772
; CURRENT APPLICATION NUMBER: US/09/609,324A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 08/920,440
; PRIOR FILING DATE: 1997-08-29
; PRIOR APPLICATION NUMBER: 08/538,709
; PRIOR FILING DATE: 1995-10-03
; PRIOR APPLICATION NUMBER: 08/318,393
; PRIOR FILING DATE: 1994-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 213
; TYPE: PRT
; ORGANISM: LERK-6
US-09-609-324A-10
```

Query Match 91.7%; Score 22; DB 1; Length 213;  
Best Local Similarity 80.0%; Pred. No. 5.1e+02;  
Matches: 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 58 GYTVE 62

## RESULT 45

US-08-920-440B-10  
; Sequence 10, Application US/08920440B  
; Patent No. 5919905  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated LERK-6  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: System 7.6  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920.440B  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2826-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 213 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-920-440B-10

Query Match 91.7%; Score 22; DB 2; Length 213;  
Best Local Similarity 80.0%; Pred. No. 5.1e+02;  
Matches: 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 58 GYTVE 62

## RESULT 46

US-09-173-492-10  
; Sequence 10, Application US/09173492  
; Patent No. 6194172  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated LERK-6  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: System 7.6  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/173,492  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,440  
; FILING DATE: 29-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2826-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 213 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-173-492-10

Query Match 91.7%; Score 22; DB 3; Length 213;  
Best Local Similarity 80.0%; Pred. No. 5.1e+02;  
Matches: 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 58 GYTVE 62

## RESULT 47

US-09-173-133-10  
; Sequence 10, Application US/09173133  
; Patent No. 6232447  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; TITLE OF INVENTION: Cytokine Designated LERK-6  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: System 7.6  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/173,133  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,440  
; FILING DATE: 29-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2826-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 213 amino acids

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-173-133-10
Query Match 91.7%; Score 22; DB 3; Length 213;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 58 GYTVE 62

RESULT 48
US-09-580-236A-10
; Sequence 10, Application US/09580236A
; Patent No. 6472174
; GENERAL INFORMATION:
; APPLICANT: Cerratti, Douglas P.
; TITLE OF INVENTION: Cytokine Designated LERK-6
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: System 7.6
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/580,236A
; FILING DATE: 26-May-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2826-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-580-236A-10
Query Match 91.7%; Score 22; DB 4; Length 213;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 58 GYTVE 62

RESULT 49
US-09-107-532A-6091
; Sequence 6091, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
```

```
;
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6091:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...232
; SEQUENCE DESCRIPTION: SEQ ID NO: 6091:
US-09-107-532A-6091
Query Match 91.7%; Score 22; DB 4; Length 232;
Best Local Similarity 80.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 67 GYTVE 71

RESULT 50
US-09-328-352-5728
; Sequence 5728, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5728
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5728
Query Match 91.7%; Score 22; DB 4; Length 250;
Best Local Similarity 80.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
```

```
Db      136 GYSVE 140
|| ||
RESULT 51
US-08-400-413-1
; Sequence 1, Application US/08400413
; Patent No. 5643737
; GENERAL INFORMATION:
; APPLICANT: Knowles Jr., Donald P.
; APPLICANT: Perryman, Lance E.
; TITLE OF INVENTION: Merozoite Proteins for Use in Detection
; TITLE OF INVENTION: of Babesia equi in Horses Using Immunological Techniques
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: USDA-ARS-OTT
; STREET: Building 005, Room 408, BARC-W
; CITY: Beltsville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20705
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,413
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/902,164
; FILING DATE: 22-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lipovsky, Joseph A.
; REGISTRATION NUMBER: 34,526
; REFERENCE/DOCKET NUMBER: 0257.92
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 504-5003
; TELEFAX: (301) 504-5060
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Babesia equi
; STRAIN: Florida
; DEVELOPMENTAL STAGE: merozoite
; IMMEDIATE SOURCE:
; CLONE: pEmail
US-08-400-413-1
Query Match      91.7%; Score 22; DB 1; Length 277;
Best Local Similarity 80.0%; Pred. No. 6.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
|| ||
Db      64 GYAVE 68

RESULT 52
US-09-270-767-40774
; Sequence 40774, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
```

```
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40774
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-40774
Query Match      91.7%; Score 22; DB 4; Length 280;
Best Local Similarity 80.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
|| ||
Db      147 GYSVE 151

RESULT 53
US-09-270-767-55990
; Sequence 55990, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55990
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-55990
Query Match      91.7%; Score 22; DB 4; Length 280;
Best Local Similarity 80.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
|| ||
Db      147 GYSVE 151

RESULT 54
US-08-972-902-7
; Sequence 7, Application US/08972902
; Patent No. 6099848
; GENERAL INFORMATION:
; APPLICANT: Frankel, Fred R.
; APPLICANT: Portnoy, Daniel A.
; TITLE OF INVENTION: BACTERIAL VACCINE VECTOR AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: One Commerce Square, 2005 Market Street, 22nd
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/972,902
; FILING DATE: 18-NOV-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Doyle-Leary, Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9596-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-972-902-7

Query Match          91.7%; Score 22; DB 3; Length 282;
Best Local Similarity 80.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      64 GYTVE 68

RESULT 55
US-09-520-207-7
; Sequence 7, Application US/09520207
; Patent No. 6504020
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Fred R.
; APPLICANT: PORTNOY, Daniel A.
; TITLE OF INVENTION: A BACTERIAL VACCINE VECTOR AND METHODS OF USE THEREOF
; FILE REFERENCE: 9596-51U1 (209596.0337)
; CURRENT APPLICATION NUMBER: US/09/520,207
; CURRENT FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: U.S. 08/972,902
; PRIOR FILING DATE: 1997-11-18
; PRIOR APPLICATION NUMBER: PCT/US98/24357
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
; US-09-520-207-7

Query Match          91.7%; Score 22; DB 4; Length 282;
Best Local Similarity 80.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      64 GYTVE 68

RESULT 56
US-10-136-253-7
; Sequence 7, Application US/10136253
; Patent No. 6635749
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Fred R.
; APPLICANT: PORTNOY, Daniel A.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACIDS COMPRISING LISTERIA DAL AND DAT GENES
; FILE REFERENCE: 053893-5011-02
; CURRENT APPLICATION NUMBER: US/10/136,253
```

```
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 09/520,207
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: PCT/US98/24357
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 08/972,902
; PRIOR FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
; US-10-136-253-7

Query Match          91.7%; Score 22; DB 4; Length 282;
Best Local Similarity 80.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      64 GYTVE 68

RESULT 57
US-08-859-106A-2
; Sequence 2, Application US/08859106A
; Patent No. 5965422
; GENERAL INFORMATION:
; APPLICANT: LOFFLER, Fridolin
; APPLICANT: NGUYEN, Quoc Khanh
; APPLICANT: SCHUSTER, Erwin
; APPLICANT: SPROBLER, Bruno
; APPLICANT: THOMAS, Lutz
; APPLICANT: WOLF, Sabine
; TITLE OF INVENTION: LYSOPHOSPHOLIPASE PRODUCED FROM
; TITLE OF INVENTION: ASPERGILLUS BY RECOMBINANT METHODS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,106A
; FILING DATE: 20-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19620649.9
; FILING DATE: 22-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 015200-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-859-106A-2

Query Match          91.7%; Score 22; DB 2; Length 297;
```



Best Local Similarity 80.0%; Pred. No. 7.2e+02; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 190 GYSVE 194

## RESULT 58

US-09-142-469-2  
; Sequence 2, Application US/09142469  
; Patent No. 6140094  
; GENERAL INFORMATION:  
; APPLICANT: LOFFLER, Fridolin  
; APPLICANT: JUNGSCHAFFER, Gerald  
; APPLICANT: KHANH, Quoc Nguyen  
; APPLICANT: SCHUSTER, Erwin  
; APPLICANT: SPROSSLER, Bruno  
; APPLICANT: WOLF, Sabine  
; TITLE OF INVENTION: PROTEIN WITH PHOSPHOLIPASE ACTIVITY  
; FILE REFERENCE: 015200-056  
; CURRENT APPLICATION NUMBER: US/09/142,469  
; CURRENT FILING DATE: 1998-09-08  
; EARLIER APPLICATION NUMBER: PCT/EP98/00081  
; EARLIER FILING DATE: 1998-01-08  
; EARLIER APPLICATION NUMBER: DE 197 01 348.1  
; EARLIER FILING DATE: 1997-01-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 297  
; TYPE: PRT  
; ORGANISM: Aspergillus lysophospholipase  
US-09-142-469-2

Query Match 91.7%; Score 22; DB 3; Length 297;

Best Local Similarity 80.0%; Pred. No. 7.2e+02; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 190 GYSVE 194

## RESULT 59

US-09-402-664A-9  
; Sequence 9, Application US/09402664A  
; Patent No. 6406723  
; GENERAL INFORMATION:  
; APPLICANT: Soe, Jorn  
; APPLICANT: Poulsen, Charlotte  
; APPLICANT: Rasmussen, Preben  
; APPLICANT: Madrid, Susan  
; APPLICANT: Zargahi, Masoud  
; TITLE OF INVENTION: IMPROVED METHOD FOR PREPARING FLOUR DOUGHS AND PRODUCTS MADE FROM  
; FILE REFERENCE: 54321.000012  
; CURRENT APPLICATION NUMBER: US/09/402,664A  
; CURRENT FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: PCT/DK98/00136  
; PRIOR FILING DATE: 1998-04-03  
; PRIOR APPLICATION NUMBER: DK 0400/97  
; PRIOR FILING DATE: 1997-04-09  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 297  
; TYPE: PRT  
; ORGANISM: Aspergillus tubingensis  
US-09-402-664A-9

Query Match 91.7%; Score 22; DB 4; Length 297;

Best Local Similarity 80.0%; Pred. No. 7.2e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 190 GYSVE 194

## RESULT 60

US-09-134-001C-3422  
; Sequence 3422, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3422  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3422

Query Match 91.7%; Score 22; DB 3; Length 300;

Best Local Similarity 80.0%; Pred. No. 7.3e+02; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 203 GYSVE 207

## RESULT 61

US-09-710-279-2122  
; Sequence 2122, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2122  
; LENGTH: 307  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-09-710-279-2122

Query Match 91.7%; Score 22; DB 4; Length 307;

Best Local Similarity 80.0%; Pred. No. 7.5e+02; Mismatches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 256 GYAVE 260

## RESULT 62

US-09-328-352-6112  
; Sequence 6112, Application US/09328352

```
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6112
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6112

Query Match          91.7%; Score 22; DB 4; Length 308;
Best Local Similarity 80.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 159 GYAVE 163

RESULT 63
US-09-252-991A-32836
; Sequence 32836, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32836
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32836

Query Match          91.7%; Score 22; DB 4; Length 313;
Best Local Similarity 80.0%; Pred. No. 7.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 75 GYSVE 79

RESULT 64
US-09-710-279-698
; Sequence 698, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 698
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-698

Query Match          91.7%; Score 22; DB 4; Length 325;
Best Local Similarity 80.0%; Pred. No. 7.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 228 GYSVE 232

RESULT 65
US-09-710-279-1330
; Sequence 1330, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1330
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1330

Query Match          91.7%; Score 22; DB 4; Length 325;
Best Local Similarity 80.0%; Pred. No. 7.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 228 GYSVE 232

RESULT 66
US-09-489-039A-10172
; Sequence 10172, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10172
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10172

Query Match          91.7%; Score 22; DB 4; Length 338;
Best Local Similarity 80.0%; Pred. No. 8.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5
Db 63 GYTVE 67
```

```
RESULT 67
US-09-134-001C-3774
; Sequence 3774, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 374
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3774

Query Match          91.7%; Score 22; DB 3; Length 345;
Best Local Similarity 80.0%; Pred. No. 8.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      ||||
Db      294 GYAVE 298

RESULT 68
US-08-762-500-77
; Sequence 77, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762,500
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,259
; FILING DATE: 17-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10469
; FILING DATE: 17-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.3
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-762-500-77

Query Match          91.7%; Score 22; DB 3; Length 349;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      ||||
Db      93 GYTVS 97

RESULT 69
US-09-252-991A-27544
; Sequence 27544, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27544
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27544

Query Match          91.7%; Score 22; DB 4; Length 349;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      ||||
Db      327 GYAVE 331

RESULT 70
US-09-107-532A-5179
; Sequence 5179, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5179:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...354
; SEQUENCE DESCRIPTION: SEQ ID NO: 5179:
US-09-107-532A-5179

Query Match          91.7%; Score 22; DB 4; Length 354;
Best Local Similarity 80.0%; Pred. No. 8.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      ||||
Db      121 GYSVE 125

RESULT 71
US-08-472-666-4
; Sequence 4, Application US/08472666
; Patent No. 5821048
; GENERAL INFORMATION:
; APPLICANT: Howley, Peter M.
; APPLICANT: Benson, John D.
; APPLICANT: Yasugi, Toshiharu
; APPLICANT: Sakai, Hiroyuki
; TITLE OF INVENTION: METHOD AND KIT FOR DIAGNOSING
; TITLE OF INVENTION: PAPILLOMAVIRUS-INFECTED CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ann-Louise Kerner, Ph.D.
; ADDRESSEE: Lappin & Kusmer
; STREET: 200 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 01209
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,666
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/677,206
; FILING DATE: 09 JUL 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: HMI-017CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-362-012A-2

Query Match          91.7%; Score 22; DB 4; Length 365;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

;
; TELEPHONE: 617-330-1300
; TELEFAX: 617-330-1311
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Human papillomavirus-16 E2
; STRAIN: HPV-16 E2
; US-08-472-666-4

Query Match          91.7%; Score 22; DB 2; Length 365;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
      ||||
Db      114 GYTVE 118

RESULT 72
US-09-362-012A-2
; Sequence 2, Application US/09362012A
; Patent No. 6432926
; GENERAL INFORMATION:
; APPLICANT: Howley, Peter M.
; APPLICANT: Dowhanick-Morrisette, Jennifer J.
; APPLICANT: Benson, John D.
; APPLICANT: Sakai, Hiroyuki
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: PAPILLOMAVIRUS-INFECTED CELLS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/362,012A
; FILING DATE: 27-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/677,206
; FILING DATE: 09 JUL 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: HMI-017CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-362-012A-2

Query Match          91.7%; Score 22; DB 4; Length 365;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 GYXVE 5  
|||  
Db 114 GYTV 118

## RESULT 73

US-09-362-012A-5  
; Sequence 5, Application US/09362012A  
; Patent No. 6432926  
; GENERAL INFORMATION:  
; APPLICANT: Dowley, Peter M.  
; ; Downhick-Morrisette, Jennifer J.  
; ; Benson, John D.  
; ; Sakai, Hiroyuki  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; ; PAPILLOMAVIRUS-INFECTED CELLS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/362,012A  
; FILING DATE: 27-Jul-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/677,206  
; FILING DATE: 09 JUL 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane E. Remillard  
; REGISTRATION NUMBER: 38,872  
; REFERENCE/DOCKET NUMBER: HMI-017CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 365 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-362-012A-5

Query Match 91.7%; Score 22; DB 4; Length 365;  
Best Local Similarity 80.0%; Pred. No. 8.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 114 GYTV 118

## RESULT 74

PCT-US96-07615-4  
; GENERAL INFORMATION:  
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE  
; TITLE OF INVENTION: METHODS, KITS, AND COMPOSITIONS FOR DIAGNOSING  
; ; PAPILLOMAVIRUS INFECTION  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lappin & Kusner  
; STREET: 200 State Street  
; CITY: Boston

; STATE: MA  
; COUNTRY: USA  
; ZIP: 01209  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/07615  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kerner, Ann-Louise  
; REGISTRATION NUMBER: 33,523  
; REFERENCE/DOCKET NUMBER: HAZ-010PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-330-1300  
; TELEFAX: 617-330-1311  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 365 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Human papillomavirus-16 E2  
; STRAIN: HPV-16 E2  
PCT-US96-07615-4

Query Match 91.7%; Score 22; DB 5; Length 365;  
Best Local Similarity 80.0%; Pred. No. 8.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 114 GYTV 118

## RESULT 75

US-09-023-023-2  
; Sequence 2, Application US/09023023  
; Patent No. 6121018  
; GENERAL INFORMATION:  
; APPLICANT: Kristine Kay Kiky  
; TITLE OF INVENTION: Interleukin-1 Beta Converting Enzyme Like Apoptosis Pr  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,023  
; FILING DATE: 12-FEB-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/042,030  
; FILING DATE: March 27, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Han, William T  
; REGISTRATION NUMBER: 34,344  
; REFERENCE/DOCKET NUMBER: GH50013  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5219  
; TELEFAX: 610-270-5090

```

;
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-023-023-2

Query Match          91.7%; Score 22; DB 3; Length 377;
Best Local Similarity 80.0%; Pred. No. 9.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
Db      170 GYTVE 174

Search completed: November 1, 2004, 21:35:58
Job time : 35 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:35:11 ; Search time 87 Seconds

(without alignments)  
22.360 Million cell updates/sec

Title: US-10-030-194A-6

Perfect score: 24

Sequence: 1 GYXVEX 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	91.7	21	US-10-611-671-46	Sequence 46, Appl
2	22	91.7	49	US-10-424-599-201461	Sequence 201461,
3	22	91.7	54	US-10-424-599-216596	Sequence 216596,
4	22	91.7	55	US-10-332-995-13	Sequence 13, Appl
5	22	91.7	60	US-10-424-599-170601	Sequence 170601,
6	22	91.7	68	US-10-138-516-7	Sequence 7, Appl
7	22	91.7	68	US-10-146-130-9	Sequence 9, Appl
8	22	91.7	68	US-10-092-934-8	Sequence 8, Appl
9	22	91.7	68	US-10-153-334-8	Sequence 8, Appl
10	22	91.7	68	US-10-198-069-8	Sequence 8, Appl
11	22	91.7	68	US-10-198-070-8	Sequence 8, Appl
12	22	91.7	68	US-10-294-891-6	Sequence 6, Appl
13	22	91.7	68	US-10-424-599-171936	Sequence 171936,

14	22	91.7	99	US-10-424-599-160975	Sequence 160975,
15	22	91.7	105	US-10-424-599-216376	Sequence 216376,
16	22	91.7	113	US-10-424-599-188662	Sequence 188662,
17	22	91.7	118	US-10-286-421-70	Sequence 70, Appl
18	22	91.7	128	US-10-767-701-55821	Sequence 55821, A
19	22	91.7	134	US-10-424-599-146245	Sequence 146245,
20	22	91.7	144	US-10-424-599-256652	Sequence 256652,
21	22	91.7	148	US-10-282-122A-60619	Sequence 60619, A
22	22	91.7	148	US-10-767-701-31757	Sequence 31757, A
23	22	91.7	151	US-10-425-114-56304	Sequence 56304, A
24	22	91.7	151	US-10-437-963-127097	Sequence 127097,
25	22	91.7	153	US-10-149-759-20	Sequence 20, Appl
26	22	91.7	153	US-10-424-599-181897	Sequence 181897,
27	22	91.7	153	US-10-424-599-225097	Sequence 225097,
28	22	91.7	155	US-10-424-599-237266	Sequence 237266,
29	22	91.7	159	US-10-767-701-62354	Sequence 62354, A
30	22	91.7	169	US-10-424-599-222627	Sequence 222627,
31	22	91.7	179	US-10-282-122A-76617	Sequence 76617, A
32	22	91.7	185	US-10-767-701-44265	Sequence 44265, A
33	22	91.7	201	US-09-738-626-6916	Sequence 6916, Ap
34	22	91.7	206	US-10-781-014-608	Sequence 608, App
35	22	91.7	209	US-09-921-984-2	Sequence 2, Appli
36	22	91.7	209	US-10-138-787-6	Sequence 6, Appli
37	22	91.7	216	US-10-767-701-35832	Sequence 35832, A
38	22	91.7	224	US-09-942-935-2	Sequence 2, Appli
39	22	91.7	238	US-10-369-493-18531	Sequence 18531, A
40	22	91.7	251	US-10-437-963-120504	Sequence 120504,
41	22	91.7	266	US-10-156-761-13406	Sequence 13406, A
42	22	91.7	271	US-10-437-963-140734	Sequence 140734,
43	22	91.7	276	US-10-424-599-173421	Sequence 173421,
44	22	91.7	279	US-10-369-493-9674	Sequence 9674, Ap
45	22	91.7	282	US-10-136-253-7	Sequence 7, Appli
46	22	91.7	296	US-10-369-493-23381	Sequence 23381, A
47	22	91.7	296	US-10-424-599-283601	Sequence 283601,
48	22	91.7	297	US-10-040-394-9	Sequence 9, Appli
49	22	91.7	297	US-10-150-429B-3	Sequence 3, Appli
50	22	91.7	297	US-10-462-527-9	Sequence 9, Appli
51	22	91.7	302	US-10-274-694-6	Sequence 6, Appli
52	22	91.7	307	US-10-282-122A-71004	Sequence 71004, A
53	22	91.7	308	US-10-282-122A-50846	Sequence 50846, A
54	22	91.7	313	US-10-369-493-122	Sequence 122, App
55	22	91.7	315	US-09-738-626-3875	Sequence 3875, Ap
56	22	91.7	321	US-10-152-886-21	Sequence 21, Appl
57	22	91.7	322	US-10-282-122A-71334	Sequence 71334, A
58	22	91.7	322	US-10-282-122A-71373	Sequence 71373, A
59	22	91.7	322	US-10-437-963-159961	Sequence 159961,
60	22	91.7	323	US-10-369-493-22968	Sequence 22968, A
61	22	91.7	325	US-09-815-242-5529	Sequence 5529, Ap
62	22	91.7	325	US-09-815-242-12354	Sequence 12354, A
63	22	91.7	325	US-10-282-122A-43904	Sequence 43904, A
64	22	91.7	325	US-10-282-122A-70689	Sequence 70689, A
65	22	91.7	328	US-10-369-493-23041	Sequence 23041, A
66	22	91.7	344	US-09-815-242-5059	Sequence 5059, Ap
67	22	91.7	344	US-10-282-122A-43511	Sequence 43511, A
68	22	91.7	344	US-10-424-599-272907	Sequence 272907,
69	22	91.7	345	US-10-369-493-21369	Sequence 21369, A
70	22	91.7	352	US-10-156-761-12213	Sequence 12213, A
71	22	91.7	355	US-10-282-122A-62715	Sequence 62715, A
72	22	91.7	359	US-10-369-493-917	Sequence 917, App
73	22	91.7	360	US-10-282-122A-74497	Sequence 74497, A
74	22	91.7	363	US-10-282-122A-60356	Sequence 60356, A
75	22	91.7	363	US-10-282-122A-147526	Sequence 147526,
76	22	91.7	365	US-10-437-963-147526	Sequence 2, Appli
77	22	91.7	365	US-10-187-536-5	Sequence 5, Appli
78	22	91.7	365	US-10-187-536-5	Sequence 1, Appli
79	22	91.7	365	US-10-047-990-3	Sequence 3, Appli
80	22	91.7	365	US-10-369-493-16467	Sequence 16467, A
81	22	91.7	365	US-10-161-927-78	Sequence 78, Appl
82	22	91.7	365	US-10-161-927-80	Sequence 80, Appl
83	22	91.7	365	US-10-425-114-48662	Sequence 48662, A
84	22	91.7	370	US-10-369-493-18602	Sequence 18602, A
85	22	91.7	373	US-10-432-443-23	Sequence 23, Appl
86	22	91.7	373	US-10-432-443-31	Sequence 31, Appl









525	22	91.7	541	14	US-10-195-891-152	Sequence 152, App	598	22	91.7	541	14	US-10-196-759-152	Sequence 152, App
526	22	91.7	541	14	US-10-196-746-152	Sequence 152, App	599	22	91.7	541	14	US-10-173-693-152	Sequence 152, App
527	22	91.7	541	14	US-10-196-752-152	Sequence 152, App	600	22	91.7	541	14	US-10-174-578-152	Sequence 152, App
528	22	91.7	541	14	US-10-196-753-152	Sequence 152, App	601	22	91.7	541	14	US-10-175-741-152	Sequence 152, App
529	22	91.7	541	14	US-10-196-761-152	Sequence 152, App	602	22	91.7	541	14	US-10-175-750-152	Sequence 152, App
530	22	91.7	541	14	US-10-197-692-152	Sequence 152, App	603	22	91.7	541	14	US-10-176-986-152	Sequence 152, App
531	22	91.7	541	14	US-10-197-693-152	Sequence 152, App	604	22	91.7	541	14	US-10-184-641-152	Sequence 152, App
532	22	91.7	541	14	US-10-197-696-152	Sequence 152, App	605	22	91.7	541	14	US-10-187-888-152	Sequence 152, App
533	22	91.7	541	14	US-10-197-698-152	Sequence 152, App	606	22	91.7	541	14	US-10-194-360-152	Sequence 152, App
534	22	91.7	541	14	US-10-197-703-152	Sequence 152, App	607	22	91.7	541	14	US-10-194-365-152	Sequence 152, App
535	22	91.7	541	14	US-10-197-711-152	Sequence 152, App	608	22	91.7	541	14	US-10-195-895-152	Sequence 152, App
536	22	91.7	541	14	US-10-198-757-152	Sequence 152, App	609	22	91.7	541	14	US-10-199-302-152	Sequence 152, App
537	22	91.7	541	14	US-10-198-761-152	Sequence 152, App	610	22	91.7	541	14	US-10-201-323-152	Sequence 152, App
538	22	91.7	541	14	US-10-198-762-152	Sequence 152, App	611	22	91.7	541	14	US-10-205-810-152	Sequence 152, App
539	22	91.7	541	14	US-10-198-763-152	Sequence 152, App	612	22	91.7	541	14	US-10-205-891-152	Sequence 152, App
540	22	91.7	541	14	US-10-198-767-152	Sequence 152, App	613	22	91.7	541	14	US-10-206-917-152	Sequence 152, App
541	22	91.7	541	14	US-10-199-301-152	Sequence 152, App	614	22	91.7	541	14	US-10-207-923-152	Sequence 152, App
542	22	91.7	541	14	US-10-199-307-152	Sequence 152, App	615	22	91.7	541	14	US-10-207-924-152	Sequence 152, App
543	22	91.7	541	14	US-10-199-312-152	Sequence 152, App	616	22	91.7	541	14	US-10-208-028-152	Sequence 152, App
544	22	91.7	541	14	US-10-199-315-152	Sequence 152, App	617	22	91.7	541	14	US-10-205-904-152	Sequence 152, App
545	22	91.7	541	14	US-10-199-316-152	Sequence 152, App	618	22	91.7	541	14	US-10-175-753-152	Sequence 152, App
546	22	91.7	541	14	US-10-199-457-152	Sequence 152, App	619	22	91.7	541	14	US-10-180-553-152	Sequence 152, App
547	22	91.7	541	14	US-10-199-459-152	Sequence 152, App	620	22	91.7	541	14	US-10-201-327-152	Sequence 152, App
548	22	91.7	541	14	US-10-199-460-152	Sequence 152, App	621	22	91.7	541	14	US-10-181-062-152	Sequence 152, App
549	22	91.7	541	14	US-10-199-461-152	Sequence 152, App	622	22	91.7	541	14	US-10-183-003-152	Sequence 152, App
550	22	91.7	541	14	US-10-199-667-152	Sequence 152, App	623	22	91.7	541	14	US-10-183-016-152	Sequence 152, App
551	22	91.7	541	14	US-10-199-673-152	Sequence 152, App	624	22	91.7	541	14	US-10-173-696-152	Sequence 152, App
552	22	91.7	541	14	US-10-201-321-152	Sequence 152, App	625	22	91.7	541	14	US-10-125-923A-152	Sequence 152, App
553	22	91.7	541	14	US-10-201-322-152	Sequence 152, App	626	22	91.7	541	14	US-10-176-491-152	Sequence 152, App
554	22	91.7	541	14	US-10-201-326-152	Sequence 152, App	627	22	91.7	541	14	US-10-176-979-152	Sequence 152, App
555	22	91.7	541	14	US-10-201-532-152	Sequence 152, App	628	22	91.7	541	14	US-10-187-592-152	Sequence 152, App
556	22	91.7	541	14	US-10-201-533-152	Sequence 152, App	629	22	91.7	541	14	US-10-197-691-152	Sequence 152, App
557	22	91.7	541	14	US-10-201-535-152	Sequence 152, App	630	22	91.7	541	14	US-10-198-771-152	Sequence 152, App
558	22	91.7	541	14	US-10-201-769-152	Sequence 152, App	631	22	91.7	541	14	US-10-174-575A-152	Sequence 152, App
559	22	91.7	541	14	US-10-201-771-152	Sequence 152, App	632	22	91.7	541	14	US-10-179-520-152	Sequence 152, App
560	22	91.7	541	14	US-10-201-854-152	Sequence 152, App	633	22	91.7	541	14	US-10-201-325-152	Sequence 152, App
561	22	91.7	541	14	US-10-202-410-152	Sequence 152, App	634	22	91.7	541	14	US-10-202-941-152	Sequence 152, App
562	22	91.7	541	14	US-10-202-473-152	Sequence 152, App	635	22	91.7	541	14	US-10-205-910-152	Sequence 152, App
563	22	91.7	541	14	US-10-202-474-152	Sequence 152, App	636	22	91.7	541	14	US-10-179-526-152	Sequence 152, App
564	22	91.7	541	14	US-10-205-503-152	Sequence 152, App	637	22	91.7	541	14	US-10-173-701-152	Sequence 152, App
565	22	91.7	541	14	US-10-205-512-152	Sequence 152, App	638	22	91.7	541	14	US-10-179-511-152	Sequence 152, App
566	22	91.7	541	14	US-10-205-521-152	Sequence 152, App	639	22	91.7	541	14	US-10-179-518-152	Sequence 152, App
567	22	91.7	541	14	US-10-208-892-152	Sequence 152, App	640	22	91.7	541	14	US-10-183-018-152	Sequence 152, App
568	22	91.7	541	14	US-10-205-896-152	Sequence 152, App	641	22	91.7	541	14	US-10-184-624-152	Sequence 152, App
569	22	91.7	541	14	US-10-205-898-152	Sequence 152, App	642	22	91.7	541	14	US-10-184-657-152	Sequence 152, App
570	22	91.7	541	14	US-10-205-901-152	Sequence 152, App	643	22	91.7	541	14	US-10-197-701-152	Sequence 152, App
571	22	91.7	541	14	US-10-205-903-152	Sequence 152, App	644	22	91.7	541	14	US-10-197-706-152	Sequence 152, App
572	22	91.7	541	14	US-10-206-909-152	Sequence 152, App	645	22	91.7	541	14	US-10-201-857-152	Sequence 152, App
573	22	91.7	541	14	US-10-206-910-152	Sequence 152, App	646	22	91.7	541	14	US-10-202-413-152	Sequence 152, App
574	22	91.7	541	14	US-10-206-911-152	Sequence 152, App	647	22	91.7	541	14	US-10-202-938-152	Sequence 152, App
575	22	91.7	541	14	US-10-206-912-152	Sequence 152, App	648	22	91.7	541	14	US-10-202-940-152	Sequence 152, App
576	22	91.7	541	14	US-10-206-913-152	Sequence 152, App	649	22	91.7	541	14	US-10-205-508-152	Sequence 152, App
577	22	91.7	541	14	US-10-206-914-152	Sequence 152, App	650	22	91.7	541	14	US-10-205-505-152	Sequence 152, App
578	22	91.7	541	14	US-10-206-920-152	Sequence 152, App	651	22	91.7	541	14	US-10-206-915-152	Sequence 152, App
579	22	91.7	541	14	US-10-206-921-152	Sequence 152, App	652	22	91.7	541	14	US-10-208-025-152	Sequence 152, App
580	22	91.7	541	14	US-10-206-923-152	Sequence 152, App	653	22	91.7	541	14	US-10-198-760-152	Sequence 152, App
581	22	91.7	541	14	US-10-206-925-152	Sequence 152, App	654	22	91.7	541	14	US-10-201-772-152	Sequence 152, App
582	22	91.7	541	14	US-10-206-926-152	Sequence 152, App	655	22	91.7	541	14	US-10-201-772-152	Sequence 152, App
583	22	91.7	541	14	US-10-206-927-152	Sequence 152, App	656	22	91.7	541	14	US-10-184-613-152	Sequence 152, App
584	22	91.7	541	14	US-10-207-916-152	Sequence 152, App	657	22	91.7	541	14	US-10-187-739-152	Sequence 152, App
585	22	91.7	541	14	US-10-207-917-152	Sequence 152, App	658	22	91.7	541	14	US-10-206-907-152	Sequence 152, App
586	22	91.7	541	14	US-10-207-918-152	Sequence 152, App	659	22	91.7	541	14	US-10-183-009-152	Sequence 152, App
587	22	91.7	541	14	US-10-207-919-152	Sequence 152, App	660	22	91.7	541	14	US-10-187-755-152	Sequence 152, App
588	22	91.7	541	14	US-10-207-920-152	Sequence 152, App	661	22	91.7	541	14	US-10-199-672-152	Sequence 152, App
589	22	91.7	541	14	US-10-207-925-152	Sequence 152, App	662	22	91.7	541	14	US-10-187-749-152	Sequence 152, App
590	22	91.7	541	14	US-10-208-021-152	Sequence 152, App	663	22	91.7	541	14	US-10-194-457-152	Sequence 152, App
591	22	91.7	541	14	US-10-208-022-152	Sequence 152, App	664	22	91.7	541	14	US-10-184-642-152	Sequence 152, App
592	22	91.7	541	14	US-10-208-023-152	Sequence 152, App	665	22	91.7	541	14	US-10-196-747-152	Sequence 152, App
593	22	91.7	541	14	US-10-208-026-152	Sequence 152, App	666	22	91.7	541	14	US-10-173-689-152	Sequence 152, App
594	22	91.7	541	14	US-10-208-029-152	Sequence 152, App	667	22	91.7	541	14	US-10-173-690-152	Sequence 152, App
595	22	91.7	541	14	US-10-208-030-152	Sequence 152, App	668	22	91.7	541	14	US-10-173-691-152	Sequence 152, App
596	22	91.7	541	14	US-10-232-232-152	Sequence 152, App	669	22	91.7	541	14	US-10-173-694-152	Sequence 152, App
597	22	91.7	541	14	US-10-195-898-152	Sequence 152, App	670	22	91.7	541	14	US-10-173-698-152	Sequence 152, App
598	22	91.7	541	14	US-10-195-899-152	Sequence 152, App	671	22	91.7	541	14	US-10-173-699-152	Sequence 152, App

671	22	91.7	541	14	US-10-173-707-152	Sequence 152, App	744	10	US-09-866-050A-626	Sequence 626, App
672	22	91.7	541	14	US-10-174-569-152	Sequence 152, App	745	14	US-10-369-493-3441	Sequence 3441, App
673	22	91.7	541	14	US-10-174-583-152	Sequence 152, App	746	10	US-09-813-153-82	Sequence 82, Appl
674	22	91.7	541	14	US-10-174-587-152	Sequence 152, App	747	9	US-09-738-626-4671	Sequence 4671, Ap
675	22	91.7	541	14	US-10-174-589-152	Sequence 152, App	748	22	US-10-158-057-205	Sequence 205, App
676	22	91.7	541	14	US-10-174-591-152	Sequence 152, App	749	22	US-10-094-749-2551	Sequence 2551, Ap
677	22	91.7	541	14	US-10-175-736-152	Sequence 152, App	750	22	US-10-282-122A-47041	Sequence 47041, A
678	22	91.7	541	14	US-10-175-742-152	Sequence 152, App	751	22	US-10-282-122A-60747	Sequence 60747, A
679	22	91.7	541	14	US-10-175-744-152	Sequence 152, App	752	22	US-10-282-122A-49935	Sequence 49935, A
680	22	91.7	541	14	US-10-175-745-152	Sequence 152, App	753	22	US-09-118-276-6	Sequence 6, Appli
681	22	91.7	541	14	US-10-175-748-152	Sequence 152, App	754	22	US-10-705-197A-6	Sequence 6, Appli
682	22	91.7	541	14	US-10-175-751-152	Sequence 152, App	755	22	US-10-437-963-139813	Sequence 129813, A
683	22	91.7	541	14	US-10-175-754-152	Sequence 152, App	756	22	US-10-389-566-1364	Sequence 1364, Ap
684	22	91.7	541	14	US-10-176-480-152	Sequence 152, App	757	22	US-10-437-963-148054	Sequence 148054, A
685	22	91.7	541	14	US-10-176-489-152	Sequence 152, App	758	22	US-10-369-493-8819	Sequence 8819, Ap
686	22	91.7	541	14	US-10-176-584-152	Sequence 152, App	759	22	US-10-369-493-4565	Sequence 4565, Ap
687	22	91.7	541	14	US-10-176-755-152	Sequence 152, App	760	22	US-10-369-493-7323	Sequence 7323, Ap
688	22	91.7	541	14	US-10-176-759-152	Sequence 152, App	761	22	US-10-369-493-107	Sequence 107, App
689	22	91.7	541	14	US-10-176-920-152	Sequence 152, App	762	22	US-10-282-122A-47771	Sequence 47771, A
690	22	91.7	541	14	US-10-176-922-152	Sequence 152, App	763	22	US-10-369-493-9757	Sequence 9757, Ap
691	22	91.7	541	14	US-10-176-924-152	Sequence 152, App	764	22	US-10-087-192-1680	Sequence 1680, Ap
692	22	91.7	541	14	US-10-176-984-152	Sequence 152, App	765	22	US-10-408-765A-2163	Sequence 2163, Ap
693	22	91.7	541	14	US-10-176-988-152	Sequence 152, App	766	22	US-10-282-122A-52202	Sequence 52202, A
694	22	91.7	541	14	US-10-179-512-152	Sequence 152, App	767	22	US-10-369-493-18955	Sequence 18955, A
695	22	91.7	541	14	US-10-179-515-152	Sequence 152, App	768	22	US-10-369-493-9308	Sequence 9308, Ap
696	22	91.7	541	14	US-10-179-592-152	Sequence 152, App	769	22	US-10-369-493-9413	Sequence 9413, Ap
697	22	91.7	541	14	US-10-173-702-152	Sequence 152, App	770	22	US-10-270-875-28	Sequence 28, Appl
698	22	91.7	541	14	US-10-173-703-152	Sequence 152, App	771	22	US-10-270-878-28	Sequence 28, Appl
699	22	91.7	541	14	US-10-173-704-152	Sequence 152, App	772	22	US-10-270-786-28	Sequence 28, Appl
700	22	91.7	541	14	US-10-174-574-152	Sequence 152, App	773	22	US-10-270-710-28	Sequence 28, Appl
701	22	91.7	541	14	US-10-176-486-152	Sequence 152, App	774	22	US-10-270-859-28	Sequence 28, Appl
702	22	91.7	541	14	US-10-176-490-152	Sequence 152, App	775	22	US-10-270-846-28	Sequence 28, Appl
703	22	91.7	541	14	US-10-176-752-152	Sequence 152, App	776	22	US-09-391-340-10	Sequence 10, Appl
704	22	91.7	541	14	US-10-176-981-152	Sequence 152, App	777	22	US-09-948-369-10	Sequence 10, Appl
705	22	91.7	541	14	US-10-176-983-152	Sequence 152, App	778	22	US-10-282-122A-59874	Sequence 59874, A
706	22	91.7	541	14	US-10-176-988-152	Sequence 152, App	779	22	US-10-094-749-2618	Sequence 2618, Ap
707	22	91.7	541	14	US-10-179-517-152	Sequence 152, App	780	22	US-10-032-583-7412	Sequence 7412, Ap
708	22	91.7	541	14	US-10-179-521-152	Sequence 152, App	781	22	US-10-264-049-2893	Sequence 2893, Ap
709	22	91.7	541	14	US-10-202-475-152	Sequence 152, App	782	22	US-10-367-095-9	Sequence 9, Appli
710	22	91.7	541	14	US-10-195-887-152	Sequence 152, App	783	22	US-10-368-046-9	Sequence 9, Appli
711	22	91.7	541	14	US-10-195-893-152	Sequence 152, App	784	22	US-10-367-367-9	Sequence 9, Appli
712	22	91.7	541	14	US-10-179-509-152	Sequence 152, App	785	22	US-10-128-714-3532	Sequence 3532, Ap
713	22	91.7	541	14	US-10-194-486-152	Sequence 152, App	786	22	US-10-282-122A-78349	Sequence 78349, A
714	22	91.7	541	14	US-10-195-900-152	Sequence 152, App	787	22	US-10-156-761-9121	Sequence 9121, Ap
715	22	91.7	541	14	US-10-198-759-152	Sequence 152, App	788	22	US-10-128-714-8532	Sequence 8532, Ap
716	22	91.7	541	14	US-10-205-506-152	Sequence 152, App	789	22	US-10-369-493-14441	Sequence 14441, A
717	22	91.7	541	14	US-10-174-570-152	Sequence 152, App	790	22	US-10-205-331-86	Sequence 86, Appl
718	22	91.7	541	14	US-10-183-005-152	Sequence 152, App	791	22	US-10-276-774-2443	Sequence 2443, Ap
719	22	91.7	541	14	US-10-179-523-152	Sequence 152, App	792	22	US-09-745-763-174	Sequence 174, App
720	22	91.7	541	14	US-10-199-463-152	Sequence 152, App	793	22	US-10-282-122A-66945	Sequence 66945, A
721	22	91.7	541	14	US-10-202-471-152	Sequence 152, App	794	22	US-10-094-749-2728	Sequence 2728, Ap
722	22	91.7	541	14	US-10-207-915-152	Sequence 152, App	795	22	US-10-094-749-2728	Sequence 2855, Ap
723	22	91.7	541	14	US-10-219-538-181	Sequence 181, App	796	22	US-10-108-260A-2855	Sequence 3079, Ap
724	22	91.7	541	14	US-10-197-709-152	Sequence 152, App	797	22	US-10-473-276-3	Sequence 3, Appli
725	22	91.7	541	15	US-10-206-915-152	Sequence 152, App	798	22	US-10-282-122A-48896	Sequence 48896, A
726	22	91.7	541	15	US-10-199-670-152	Sequence 152, App	799	22	US-10-369-493-23117	Sequence 23117, A
727	22	91.7	541	15	US-10-201-858-152	Sequence 152, App	800	22	US-10-389-566-2409	Sequence 2409, Ap
728	22	91.7	541	15	US-10-205-890-152	Sequence 152, App	801	22	US-10-416-330-24	Sequence 24, Appl
729	22	91.7	541	15	US-10-208-024-152	Sequence 152, App	802	22	US-10-094-749-3028	Sequence 3028, Ap
730	22	91.7	541	15	US-10-201-853-152	Sequence 152, App	803	22	US-10-369-493-11460	Sequence 11460, A
731	22	91.7	541	15	US-10-206-916-152	Sequence 152, App	804	22	US-10-369-493-14726	Sequence 14726, A
732	22	91.7	545	9	US-09-815-242-5591	Sequence 5591, Ap	805	22	US-10-369-493-15207	Sequence 15207, A
733	22	91.7	550	17	US-10-796-953-14	Sequence 14, Appl	806	22	US-10-437-963-152562	Sequence 152562, A
734	22	91.7	553	9	US-09-815-242-12198	Sequence 12198, A	807	22	US-10-437-963-20447	Sequence 20447, A
735	22	91.7	553	15	US-10-282-122A-44061	Sequence 44061, A	808	22	US-10-369-493-20447	Sequence 20447, A
736	22	91.7	555	14	US-10-094-749-2279	Sequence 2279, Ap	809	22	US-10-341-434-141	Sequence 141, App
737	22	91.7	557	15	US-10-282-122A-58498	Sequence 58498, A	810	22	US-10-408-765A-554	Sequence 554, App
738	22	91.7	558	16	US-10-437-963-181571	Sequence 181571, A	811	22	US-09-808-602-67	Sequence 67, Appl
739	22	91.7	561	15	US-10-264-237-2674	Sequence 2674, Ap	812	22	US-09-800-198-56	Sequence 56, Appl
740	22	91.7	565	15	US-10-627-476-616	Sequence 616, App	813	22	US-10-289-776-15	Sequence 15, Appl
741	22	91.7	568	16	US-10-437-963-105035	Sequence 105035, A	814	22	US-10-282-122A-46029	Sequence 46029, A
742	22	91.7	572	16	US-10-437-963-126711	Sequence 126711, A	815	22	US-10-389-566-2154	Sequence 2154, Ap
743	22	91.7	575	16	US-10-437-963-173223	Sequence 173223, A	816	22	US-10-437-963-175778	Sequence 175778, A

817	22	91.7	1659	9	US-09-801-368-118	Sequence 118, App	890	21	87.5	83	15	US-10-424-599-230404	Sequence 230404,
818	22	91.7	1664	16	US-10-437-963-175780	Sequence 175780, App	891	21	87.5	83	16	US-10-437-963-133769	Sequence 133769,
819	22	91.7	1752	14	US-10-116-275-205	Sequence 205, App	892	21	87.5	84	9	US-09-864-761-41934	Sequence 41934, A
820	22	91.7	1752	14	US-10-295-027-360	Sequence 360, App	893	21	87.5	85	15	US-10-424-599-153988	Sequence 153988,
821	22	91.7	1787	14	US-10-369-493-1159	Sequence 1159, App	894	21	87.5	85	15	US-10-424-599-238629	Sequence 238629,
822	22	91.7	1805	16	US-10-408-765A-274	Sequence 274, App	895	21	87.5	86	14	US-10-038-010-16	Sequence 16, Appl
823	22	91.7	1822	14	US-10-295-027-1254	Sequence 1254, App	896	21	87.5	86	15	US-10-087-684-96	Sequence 96, Appl
824	22	91.7	1873	14	US-10-392-113-6	Sequence 6, Appl	897	21	87.5	86	15	US-10-218-779-96	Sequence 96, Appl
825	22	91.7	1873	14	US-10-392-113-17	Sequence 17, Appl	898	21	87.5	86	16	US-10-437-963-172409	Sequence 172409,
826	22	91.7	2166	16	US-10-437-963-118281	Sequence 118281, App	899	21	87.5	87	16	US-10-767-701-40109	Sequence 40109, A
827	22	91.7	2905	16	US-10-437-963-154118	Sequence 154118, App	900	21	87.5	88	14	US-10-029-386-33979	Sequence 33979, A
828	22	91.7	5701	9	US-09-864-761-37319	Sequence 37319, A	901	21	87.5	88	16	US-10-437-963-167312	Sequence 167312,
829	22	91.7	5701	14	US-10-029-386-32438	Sequence 32438, A	902	21	87.5	89	15	US-10-424-599-166708	Sequence 166708,
830	22	91.7	26926	9	US-09-759-508B-2	Sequence 2, Appl	903	21	87.5	90	9	US-09-864-761-42006	Sequence 42006, A
831	21	87.5	9	14	US-10-245-871-700	Sequence 700, App	904	21	87.5	91	16	US-10-767-701-53455	Sequence 53455, A
832	21	87.5	9	15	US-10-253-286-700	Sequence 700, App	905	21	87.5	92	9	US-09-993-844-21	Sequence 21, Appl
833	21	87.5	10	10	US-09-572-404B-2917	Sequence 2917, App	906	21	87.5	92	15	US-10-633-438-7	Sequence 7, Appl
834	21	87.5	11	9	US-09-748-038B-1	Sequence 1, Appl	907	21	87.5	94	16	US-10-240-430-9	Sequence 9, Appl
835	21	87.5	11	9	US-09-748-038B-2	Sequence 2, Appl	908	21	87.5	95	9	US-09-887-593-8	Sequence 8, Appl
836	21	87.5	12	8	US-08-736-019-164	Sequence 164, App	909	21	87.5	95	15	US-10-424-599-162499	Sequence 162499,
837	21	87.5	18	14	US-10-225-567A-1814	Sequence 814, App	910	21	87.5	95	15	US-10-335-977-8299	Sequence 8299, App
838	21	87.5	21	9	US-09-826-069A-16	Sequence 16, Appl	911	21	87.5	98	9	US-09-864-761-34731	Sequence 34731, A
839	21	87.5	21	9	US-09-826-069A-17	Sequence 17, Appl	912	21	87.5	100	14	US-10-104-047-2362	Sequence 2362, App
840	21	87.5	21	10	US-09-988-165B-11	Sequence 11, Appl	913	21	87.5	101	9	US-09-864-761-38344	Sequence 38344, A
841	21	87.5	26	14	US-10-245-871-723	Sequence 723, App	914	21	87.5	101	14	US-10-029-386-28997	Sequence 28997, A
842	21	87.5	26	15	US-10-253-286-723	Sequence 723, App	915	21	87.5	102	9	US-09-864-761-38170	Sequence 38170, A
843	21	87.5	29	15	US-10-181-937-11	Sequence 11, Appl	916	21	87.5	102	10	US-09-933-261-8	Sequence 8, Appl
844	21	87.5	32	14	US-10-179-382-11	Sequence 11, Appl	917	21	87.5	103	14	US-10-256-702-8	Sequence 8, Appl
845	21	87.5	35	15	US-10-424-599-150399	Sequence 150399, App	918	21	87.5	103	16	US-10-229-567-39	Sequence 39, Appl
846	21	87.5	37	9	US-09-864-761-44411	Sequence 44411, A	919	21	87.5	103	16	US-10-437-963-178878	Sequence 178878,
847	21	87.5	37	16	US-10-437-963-177100	Sequence 177100, App	920	21	87.5	104	15	US-10-218-779-97	Sequence 97, Appl
848	21	87.5	38	15	US-10-424-599-263924	Sequence 263924, App	921	21	87.5	105	15	US-10-424-599-223426	Sequence 223426,
849	21	87.5	44	16	US-10-437-963-188107	Sequence 188107, App	922	21	87.5	106	14	US-10-369-493-1185	Sequence 1185, App
850	21	87.5	47	14	US-10-106-698-6108	Sequence 6108, App	923	21	87.5	106	15	US-10-424-599-144420	Sequence 144420,
851	21	87.5	51	10	US-09-988-165B-14	Sequence 14, Appl	924	21	87.5	106	15	US-10-424-599-265466	Sequence 265466,
852	21	87.5	51	15	US-10-424-599-149163	Sequence 149163, App	925	21	87.5	107	15	US-10-424-599-161705	Sequence 161705,
853	21	87.5	51	15	US-10-424-599-240623	Sequence 240623, App	926	21	87.5	107	15	US-10-424-599-209415	Sequence 209415,
854	21	87.5	51	15	US-10-424-599-258112	Sequence 258112, App	927	21	87.5	107	15	US-10-424-599-268557	Sequence 268557,
855	21	87.5	51	16	US-10-437-963-147144	Sequence 147144, App	928	21	87.5	107	15	US-10-425-114-62277	Sequence 62277, A
856	21	87.5	53	15	US-10-424-599-249850	Sequence 249850, App	929	21	87.5	107	16	US-10-437-963-103692	Sequence 103692,
857	21	87.5	53	15	US-10-424-599-272287	Sequence 272287, App	930	21	87.5	108	14	US-10-156-761-7658	Sequence 7658, App
858	21	87.5	55	16	US-10-437-963-185341	Sequence 185341, App	931	21	87.5	110	9	US-09-925-300-1328	Sequence 1328, App
859	21	87.5	57	15	US-10-424-599-164827	Sequence 164827, App	932	21	87.5	110	11	US-09-964-956-69	Sequence 69, App
860	21	87.5	58	9	US-09-738-626-3949	Sequence 3949, App	933	21	87.5	110	14	US-10-156-761-8628	Sequence 8628, App
861	21	87.5	58	15	US-10-424-599-225080	Sequence 225080, App	934	21	87.5	110	15	US-10-072-012-806	Sequence 806, App
862	21	87.5	58	15	US-10-424-599-246711	Sequence 246711, App	935	21	87.5	111	16	US-10-767-701-58766	Sequence 58766, A
863	21	87.5	59	15	US-10-424-599-237501	Sequence 237501, App	936	21	87.5	112	11	US-09-864-408A-3458	Sequence 3458, App
864	21	87.5	59	15	US-10-424-599-251262	Sequence 251262, App	937	21	87.5	112	14	US-10-194-975-116	Sequence 116, App
865	21	87.5	60	15	US-10-424-599-156574	Sequence 156574, App	938	21	87.5	114	16	US-10-767-701-32083	Sequence 32083, A
866	21	87.5	62	10	US-09-988-165B-15	Sequence 15, Appl	939	21	87.5	115	10	US-09-764-891-5131	Sequence 5131, App
867	21	87.5	62	14	US-10-029-386-31738	Sequence 31738, A	940	21	87.5	115	13	US-10-027-450-13	Sequence 13, Appl
868	21	87.5	62	15	US-10-131-487A-123	Sequence 123, App	941	21	87.5	116	14	US-10-229-567-38	Sequence 38, Appl
869	21	87.5	63	15	US-10-424-599-195480	Sequence 195480, App	942	21	87.5	117	14	US-10-156-761-8280	Sequence 8280, App
870	21	87.5	64	14	US-10-264-480-24	Sequence 24, Appl	943	21	87.5	118	14	US-10-410-681-20	Sequence 20, Appl
871	21	87.5	65	15	US-10-282-122A-51889	Sequence 51889, App	944	21	87.5	118	15	US-10-424-599-155357	Sequence 155357,
872	21	87.5	66	14	US-10-264-480-21	Sequence 21, Appl	945	21	87.5	118	15	US-10-335-977-7922	Sequence 7922, App
873	21	87.5	66	15	US-10-424-599-231248	Sequence 231248, App	946	21	87.5	119	15	US-10-424-599-163523	Sequence 163523,
874	21	87.5	67	15	US-10-424-599-212021	Sequence 212021, App	947	21	87.5	119	15	US-10-424-599-167052	Sequence 167052,
875	21	87.5	67	15	US-10-437-963-151127	Sequence 151127, App	948	21	87.5	119	15	US-10-424-599-174434	Sequence 174434,
876	21	87.5	69	13	US-10-001-879-134	Sequence 134, App	949	21	87.5	119	15	US-10-425-114-44265	Sequence 44265, A
877	21	87.5	69	15	US-10-424-599-159390	Sequence 159390, App	950	21	87.5	119	15	US-10-335-977-8300	Sequence 8300, App
878	21	87.5	70	16	US-10-437-963-129606	Sequence 129606, App	951	21	87.5	119	16	US-10-437-963-154860	Sequence 154860,
879	21	87.5	72	16	US-10-437-963-118221	Sequence 118221, App	952	21	87.5	120	15	US-10-335-977-7921	Sequence 7921, App
880	21	87.5	73	9	US-09-993-844-63	Sequence 63, Appl	953	21	87.5	121	9	US-09-815-242-5545	Sequence 5545, App
881	21	87.5	74	15	US-10-424-599-191602	Sequence 191602, App	954	21	87.5	121	9	US-09-815-242-12229	Sequence 12229, A
882	21	87.5	75	16	US-10-437-963-120918	Sequence 120918, App	955	21	87.5	121	9	US-08-815-242-12836	Sequence 12836, A
883	21	87.5	76	15	US-10-424-599-198290	Sequence 198290, App	956	21	87.5	121	9	US-09-908-931B-40	Sequence 40, Appl
884	21	87.5	76	16	US-10-437-963-118997	Sequence 118997, App	957	21	87.5	121	15	US-10-282-122A-44178	Sequence 44178, A
885	21	87.5	77	15	US-10-424-599-194582	Sequence 194582, App	958	21	87.5	121	15	US-10-282-122A-70530	Sequence 70530, A
886	21	87.5	79	16	US-10-437-963-149206	Sequence 149206, App	959	21	87.5	121	15	US-10-332-964-41	Sequence 41, Appl
887	21	87.5	79	16	US-10-437-963-160107	Sequence 160107, App	960	21	87.5	121	16	US-10-437-963-127109	Sequence 127109,
888	21	87.5	80	15	US-10-087-684-97	Sequence 97, Appl	961	21	87.5	122	15	US-10-424-599-160224	Sequence 160224,
889	21	87.5	83	9	US-09-993-844-64	Sequence 64, Appl	962	21	87.5	122	15	US-10-424-599-167407	Sequence 167407,

963 21 87.5 122 16 US-10-767-701-63030 Sequence 63030, A  
964 21 87.5 123 15 US-10-424-599-272017 Sequence 272017, A  
965 21 87.5 125 16 US-10-437-963-149639 Sequence 149639, A  
966 21 87.5 127 16 US-10-767-701-43194 Sequence 43194, A  
967 21 87.5 130 15 US-10-424-599-204329 Sequence 204329, A  
968 21 87.5 130 15 US-10-424-599-264852 Sequence 264852, A  
969 21 87.5 130 15 US-10-425-114-49862 Sequence 49862, A  
970 21 87.5 131 15 US-10-296-115-1231 Sequence 1231, Ap  
971 21 87.5 131 16 US-10-437-963-203799 Sequence 203799, A  
972 21 87.5 131 16 US-10-767-701-59858 Sequence 59858, A  
973 21 87.5 132 15 US-10-424-599-145382 Sequence 145382, A  
974 21 87.5 132 15 US-10-424-599-159938 Sequence 159938, A  
975 21 87.5 134 15 US-10-424-599-156761 Sequence 156761, A  
976 21 87.5 134 15 US-10-424-599-270227 Sequence 270227, A  
977 21 87.5 135 16 US-10-767-701-59112 Sequence 59112, A  
978 21 87.5 136 9 US-09-864-761-38243 Sequence 38243, A  
979 21 87.5 136 14 US-10-029-386-32312 Sequence 32312, A  
980 21 87.5 137 16 US-10-767-701-48877 Sequence 48877, A  
981 21 87.5 138 15 US-10-424-599-187030 Sequence 187030, A  
982 21 87.5 138 16 US-10-437-963-185069 Sequence 185069, A  
983 21 87.5 139 9 US-09-864-761-34885 Sequence 34885, A  
984 21 87.5 139 15 US-10-424-599-245235 Sequence 245235, A  
985 21 87.5 140 14 US-10-038-010-18 Sequence 18, Appl  
986 21 87.5 140 14 US-10-332-282-15 Sequence 15, Appl  
987 21 87.5 142 11 US-09-864-408A-8010 Sequence 8010, Ap  
988 21 87.5 142 14 US-10-029-386-30849 Sequence 30849, A  
989 21 87.5 142 15 US-10-442-502-18 Sequence 18, Appl  
990 21 87.5 142 15 US-10-424-599-225610 Sequence 225610, A  
991 21 87.5 143 15 US-10-282-122A-73929 Sequence 73929, A  
992 21 87.5 143 15 US-10-424-599-169951 Sequence 169951, A  
993 21 87.5 144 13 US-10-062-254-34 Sequence 34, Appl  
994 21 87.5 144 14 US-10-029-386-28819 Sequence 28819, A  
995 21 87.5 145 15 US-10-424-599-216365 Sequence 216365, A  
996 21 87.5 145 16 US-10-437-963-146357 Sequence 146357, A  
997 21 87.5 145 16 US-10-437-963-146368 Sequence 146368, A  
998 21 87.5 146 15 US-10-424-599-198169 Sequence 198169, A  
999 21 87.5 146 16 US-10-767-701-40676 Sequence 40676, A  
1000 21 87.5 147 15 US-10-424-599-172950 Sequence 172950, A

ALIGNMENTS

RESULT 1  
US-10-611-671-46  
; Sequence 46, Application US/10611671  
; Publication No. US20040091907A1  
; GENERAL INFORMATION:  
; APPLICANT: SWEAL, TOD R.  
; APPLICANT: CALLOW, MARINELLA G.  
; APPLICANT: JALLAL, BAHJA  
; APPLICANT: ZOZULYA, SERGEY  
; APPLICANT: GISHIZKY, MIKHAIL L.  
; TITLE OF INVENTION: GEF-HIB: BIOMARKERS, COMPLEXES, ASSAYS AND THERAPEUTIC  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 034536-0407  
; CURRENT APPLICATION NUMBER: US/10/611,671  
; PRIOR FILING DATE: 2003-07-02  
; PRIOR APPLICATION NUMBER: 60/460,053  
; PRIOR FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: 60/393,600  
; PRIOR FILING DATE: 2002-07-05  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 46  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-611-671-46

Query Match 91.7%; Score 22; DB 15; Length 21;  
Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 GYXVE 5  
|||  
Db 16 GYSVE 20  
  
RESULT 2  
US-10-424-599-201461  
; Sequence 201461, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 201461  
; LENGTH: 49  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_23945C.1.pep  
US-10-424-599-201461

Query Match 91.7%; Score 22; DB 15; Length 49;  
Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 GYXVE 5  
|||  
Db 17 GYSVE 21

RESULT 3  
US-10-424-599-216596  
; Sequence 216596, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 216596  
; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37614C.1.pep  
US-10-424-599-216596

Query Match 91.7%; Score 22; DB 15; Length 54;  
Best Local Similarity 80.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 GYXVE 5  
|||  
Db 29 GYSVE 33

```
RESULT 4
US-10-332-995-13
; Sequence 13, Application US/10332995
; Publication No. US2004004349A1
; GENERAL INFORMATION:
; APPLICANT: National Research Council of Canada
; APPLICANT: Zou, Jitao
; APPLICANT: Zheng, Zhifu
; TITLE OF INVENTION: Glycylol-3-Phosphate/Dihydroxyacetone Phosphate Dual Substrate
; TITLE OF INVENTION: Acyltransferases
; FILE REFERENCE: 46617-PT
; CURRENT APPLICATION NUMBER: US/10/332,995
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: US60/220,482
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-332-995-13

Query Match          91.7%; Score 22; DB 15; Length 55;
Best Local Similarity 80.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
   ||||
Db 6 GYSVE 10

RESULT 5
US-10-424-599-170601
; Sequence 170601, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170601
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125068C.1.pep
US-10-424-599-170601

Query Match          91.7%; Score 22; DB 15; Length 60;
Best Local Similarity 80.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
   ||||
Db 6 GYSVE 10

RESULT 6
US-10-138-516-7
; Sequence 7, Application US/10138516
; Publication No. US20030003445A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHOD OF PREVENTING CELL DEATH USING ANTIBODIES TO
; TITLE OF INVENTION: NEURAL THREAD PROTEINS
; FILE REFERENCE: 59003.000004
```

```
; CURRENT APPLICATION NUMBER: US/10/138,516
; CURRENT FILING DATE: 2002-07-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-138-516-7

Query Match          91.7%; Score 22; DB 14; Length 68;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
   ||||
Db 43 GYAVE 47

RESULT 7
US-10-146-130-9
; Sequence 9, Application US/10146130
; Publication No. US20030004107A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHOD OF PREVENTING CELL DEATH USING SEGMENTS OF
; TITLE OF INVENTION: NEURAL THREAD PROTEINS
; FILE REFERENCE: 59003.000007
; CURRENT APPLICATION NUMBER: US/10/146,130
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-130-9

Query Match          91.7%; Score 22; DB 14; Length 68;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
   ||||
Db 43 GYAVE 47

RESULT 8
US-10-092-934-8
; Sequence 8, Application US/10092934
; Publication No. US2003005490A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHODS OF USING NEURAL THREAD PROTEINS TO TREAT TUMORS
; TITLE OF INVENTION: AND CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 018792-0199
; CURRENT APPLICATION NUMBER: US/10/092,934
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/273,957
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Neural thread
; OTHER INFORMATION: protein
US-10-092-934-8

Query Match          91.7%; Score 22; DB 14; Length 68;
```

Best Local Similarity 80.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 43 GYAVE 47

RESULT 9  
US-10-153-334-8  
; Sequence 8, Application US/10153334  
; Publication No. US20030096350A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: 59003.000006  
; CURRENT APPLICATION NUMBER: US/10/153,334  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,156  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-153-334-8

Query Match 91.7%; Score 22; DB 14; Length 68;  
Best Local Similarity 80.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 43 GYAVE 47

RESULT 10  
US-10-198-069-8  
; Sequence 8, Application US/10198069  
; Publication No. US20030096756A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: 59003.000009  
; CURRENT APPLICATION NUMBER: US/10/198,069  
; CURRENT FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 60/306,161  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/306,150  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/331,477  
; PRIOR FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Unknown NTP  
; OTHER INFORMATION: peptide  
US-10-198-069-8

Query Match 91.7%; Score 22; DB 14; Length 68;  
Best Local Similarity 80.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 43 GYAVE 47

RESULT 11  
US-10-198-070-8  
; Sequence 8, Application US/10198070  
; Publication No. US20030109437A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; APPLICANT: GEMMEL, JACK  
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: 59003.000008  
; CURRENT APPLICATION NUMBER: US/10/198,070  
; CURRENT FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 60/306,161  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/306,150  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/331,477  
; PRIOR FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 125  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Unknown NTP  
; OTHER INFORMATION: peptide  
US-10-198-070-8

Query Match 91.7%; Score 22; DB 14; Length 68;  
Best Local Similarity 80.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 43 GYAVE 47

RESULT 12  
US-10-294-891-6  
; Sequence 6, Application US/10294891  
; Publication No. US20030166569A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; APPLICANT: GEMMEL, JACK  
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF CELLS  
; FILE REFERENCE: 59003.000010  
; CURRENT APPLICATION NUMBER: US/10/294,891  
; CURRENT FILING DATE: 2002-11-15  
; PRIOR APPLICATION NUMBER: 60/331,447  
; PRIOR FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 6  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-294-891-6

Query Match 91.7%; Score 22; DB 14; Length 68;  
Best Local Similarity 80.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 43 GYAVE 47



## RESULT 13

US-10-424-599-171936  
; Sequence 171936, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 171936  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: Glycine max  
; LOCATION: (1)..(105)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_126272C.1.pep  
US-10-424-599-171936

Query Match 91.7%; Score 22; DB 15; Length 68;  
Best Local Similarity 80.0%; Pred. No. 6.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 20 GYSVE 24

## RESULT 14

US-10-424-599-160975  
; Sequence 160975, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 160975  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(99)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_116378C.1.pep  
US-10-424-599-160975

Query Match 91.7%; Score 22; DB 15; Length 99;  
Best Local Similarity 80.0%; Pred. No. 9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 61 GYAVE 65

## RESULT 15

US-10-424-599-216376  
; Sequence 216376, Application US/10424599  
; Publication No. US20040031072A1

; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 216376  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(105)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37416C.1.pep  
US-10-424-599-216376

Query Match 91.7%; Score 22; DB 15; Length 105;  
Best Local Similarity 80.0%; Pred. No. 9.5e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 47 GYSVE 51

## RESULT 16

US-10-424-599-188662  
; Sequence 188662, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 188662  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(113)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_141376C.1.pep  
US-10-424-599-188662

Query Match 91.7%; Score 22; DB 15; Length 113;  
Best Local Similarity 80.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
Db 76 GYAVE 80

## RESULT 17

US-10-286-421-70  
; Sequence 70, Application US/10286421  
; Publication No. US20030129656A1  
; GENERAL INFORMATION:

; APPLICANT: Park, Frances  
; APPLICANT: Gajiwala, Ketan S.  
; APPLICANT: Buchanan, Sean Grant  
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF BACTERIAL  
; TITLE OF INVENTION: NUCLEIC ACID BINDING PROTEIN  
; FILE REFERENCE: 524982001000  
; CURRENT APPLICATION NUMBER: US/10/286,421  
; CURRENT FILING DATE: 2003-02-11  
; PRIOR APPLICATION NUMBER: 60/336,899  
; PRIOR FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 70  
; LENGTH: 118  
; TYPE: PRT  
; ORGANISM: Deinococcus radiodurans  
US-10-286-421-70

Query Match 91.7%; Score 22; DB 14; Length 118;  
Best Local Similarity 80.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 82 GYTVE 86

## RESULT 18

US-10-767-701-55821  
; Sequence 55821, Application US/107677701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 55821  
; LENGTH: 128  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 30933070.pep  
US-10-767-701-55821

Query Match 91.7%; Score 22; DB 16; Length 128;  
Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 64 GYAVE 68

## RESULT 19

US-10-424-599-146245  
; Sequence 146245, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 146245  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(134)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_103077C.1.pep  
US-10-424-599-146245

Query Match 91.7%; Score 22; DB 15; Length 134;  
Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 122 GYSVE 126

## RESULT 20

US-10-424-599-256652  
; Sequence 256652, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 256652  
; LENGTH: 144  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(144)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_73780C.1.pep  
US-10-424-599-256652

Query Match 91.7%; Score 22; DB 15; Length 144;  
Best Local Similarity 80.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 41 GYTVE 45

## RESULT 21

US-10-282-122A-60619  
; Sequence 60619, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.

APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 60619  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: Listeria monocytogenes  
US-10-282-122A-60619

Query Match 91.7%; Score 22; DB 15; Length 148;  
Best Local Similarity 80.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 34 GYAVE 38

RESULT 22  
US-10-767-701-31757  
; Sequence 31757, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 31757  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C10451\_1.pep  
US-10-767-701-31757

Query Match 91.7%; Score 22; DB 16; Length 148;  
Best Local Similarity 80.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 116 GYSVE 120

RESULT 23  
US-10-425-114-56304  
; Sequence 56304, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 56304  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700975650\_FLI.pep  
US-10-425-114-56304

Query Match 91.7%; Score 22; DB 15; Length 151;  
Best Local Similarity 80.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 111 GYSVE 115

RESULT 24  
US-10-437-963-127097  
; Sequence 127097, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 127097  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_29581C.1.pep  
US-10-437-963-127097

Query Match 91.7%; Score 22; DB 16; Length 151;  
Best Local Similarity 80.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 46 GYAVE 50

RESULT 25  
US-10-149-759-20

```
; Sequence 20, Application US/10149759
; Publication No. US20030157592A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirpus, Petra
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins
; TITLE OF INVENTION: involved in the synthesis of tocopherols and
; TITLE OF INVENTION: carotenoids.
; FILE REFERENCE: BASF/NAE 133/99 PCT/US
; CURRENT APPLICATION NUMBER: US/10/149,759
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: PCT/EP/00/12698
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: WordPerfect version 6.1
; SEQ ID NO 20
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-10-149-759-20

Query Match          91.7%; Score 22; DB 14; Length 153;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 140 GYTVE 144

RESULT 26
US-10-424-599-181897
; Sequence 181897, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 181897
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(155)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135267C.1.pep
US-10-424-599-181897

Query Match          91.7%; Score 22; DB 15; Length 153;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 33 GYTVE 37

RESULT 27
US-10-424-599-225097
; Sequence 225097, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 225097
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(155)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45293C.1.pep
US-10-424-599-225097

Query Match          91.7%; Score 22; DB 15; Length 153;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 68 GYTVE 72

RESULT 28
US-10-424-599-237266
; Sequence 237266, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 237266
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(155)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_56276C.1.pep
US-10-424-599-237266

Query Match          91.7%; Score 22; DB 15; Length 155;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
Db 5 GYAVE 9

RESULT 29
US-10-767-701-62354
; Sequence 62354, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
```

```
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 62354
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 18059887.pep
US-10-767-701-62354

Query Match          91.7%; Score 22; DB 16; Length 159;
Best Local Similarity 80.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
   ||||
Db 89 GYSVE 93

RESULT 30
US-10-424-599-222627
; Sequence 222627, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 222627
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_43060C.1.pep
US-10-424-599-222627

Query Match          91.7%; Score 22; DB 15; Length 169;
Best Local Similarity 80.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
   ||||
Db 42 GYAVE 46

RESULT 31
US-10-282-122A-76617
; Sequence 76617, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76617
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Treponema pallidum
US-10-282-122A-76617

Query Match          91.7%; Score 22; DB 15; Length 179;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
   ||||
Db 155 GYSVE 159

RESULT 32
US-10-767-701-44265
; Sequence 44265, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 44265
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C7224_1.pep
US-10-767-701-44265

Query Match          91.7%; Score 22; DB 16; Length 185;
Best Local Similarity 80.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5
   ||||
Db 157 GYAVE 161

RESULT 33
US-09-738-626-6916
; Sequence 6916, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
```

```
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6916
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6916

Query Match          91.7%; Score 22; DB 9; Length 201;
Best Local Similarity 80.0%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
Db      152 GYTV 156

RESULT 34
US-10-781-014-608
; Sequence 608, Application US/10781014
; Publication No. US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
; FILE REFERENCE: BGI-126CPCN
; CURRENT APPLICATION NUMBER: US/10/781,014
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 09/602,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/143,208
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/151,572
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19931412.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931424.1
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931428.4
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 784
; SEQ ID NO 608
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-781-014-608

Query Match          91.7%; Score 22; DB 16; Length 206;
Best Local Similarity 80.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
Db      37 GYAVE 41

RESULT 35
US-09-921-984-2
; Sequence 2, Application US/09921984
; Patent No. US20020156239A1
; GENERAL INFORMATION:
; APPLICANT: Flanagan, John G.
; TITLE OF INVENTION: EPF Receptor Ligands, and Uses Related Thereto
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(txt)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/921,984
; FILING DATE: 31-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/308,814
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-921-984-2

Query Match          91.7%; Score 22; DB 9; Length 209;
Best Local Similarity 80.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
Db      54 GYTV 58

RESULT 36
US-10-138-787-6
; Sequence 6, Application US/10138787
```

Publication No. US20020172984A1  
; GENERAL INFORMATION:  
; APPLICANT: Holland, Sacha  
; APPLICANT: Mbamalu, Geraldine  
; APPLICANT: Pawson, Tony  
; TITLE OF INVENTION: OLIGOMERIZED RECEPTORS WHICH AFFECT PATHWAYS REGULATED  
; TITLE OF INVENTION: BY TRANSMEMBRANE LIGANDS FOR ELK-RELATED RECEPTOR  
; TITLE OF INVENTION: TYROSINE KINASES  
; FILE REFERENCE: 11757.23USWO  
; CURRENT APPLICATION NUMBER: US/10/138,787  
; CURRENT FILING DATE: 2002-05-03  
; PRIOR APPLICATION NUMBER: US/09/214,631  
; PRIOR FILING DATE: 1999-03-12  
; PRIOR APPLICATION NUMBER: PCT/CA97/00473  
; PRIOR FILING DATE: 1997-07-04  
; PRIOR APPLICATION NUMBER: 60/021,272  
; PRIOR FILING DATE: 1996-07-05  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-138-787-6

Query Match 91.7%; Score 22; DB 13; Length 209;  
Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 54 GYTV 58

RESULT 37  
US-10-767-701-35832  
; Sequence 35832, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 35832  
; LENGTH: 216  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(216)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C75823\_1.pep  
US-10-767-701-35832

Query Match 91.7%; Score 22; DB 16; Length 216;  
Best Local Similarity 80.0%; Pred. No. 2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 112 GYAVE 116

RESULT 38  
US-09-942-935-2  
; Sequence 2, Application US/09942935  
; Patent No. US2002010675A1

; GENERAL INFORMATION:  
; APPLICANT: Degussa AG  
; TITLE OF INVENTION: Nucleotide Sequences Coding for the sigM Gene  
; FILE REFERENCE: 000449 BT  
; CURRENT APPLICATION NUMBER: US/09/942,935  
; CURRENT FILING DATE: 2001-12-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-942-935-2

Query Match 91.7%; Score 22; DB 9; Length 224;  
Best Local Similarity 80.0%; Pred. No. 2.1e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 175 GYTV 179

RESULT 39  
US-10-369-493-18531  
; Sequence 18531, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 18531  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Halobacterium sp. NRC-1  
US-10-369-493-18531

Query Match 91.7%; Score 22; DB 14; Length 238;  
Best Local Similarity 80.0%; Pred. No. 2.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 206 GYTV 210

RESULT 40  
US-10-437-963-120504  
; Sequence 120504, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 120504  
; LENGTH: 251  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_23617C.1.pep  
US-10-437-963-120504

Query Match 91.7%; Score 22; DB 16; Length 251;  
Best Local Similarity 80.0%; Pred. No. 2.3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 5 GYVE 9

RESULT 41  
US-10-156-761-13406  
; Sequence 13406, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 13406  
; LENGTH: 266  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-13406

Query Match 91.7%; Score 22; DB 14; Length 266;  
Best Local Similarity 80.0%; Pred. No. 2.5e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 241 GYAVE 245

RESULT 42  
US-10-437-963-140734  
; Sequence 140734, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 140734  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_41905C.1.pep  
US-10-437-963-140734

Query Match 91.7%; Score 22; DB 16; Length 271;  
Best Local Similarity 80.0%; Pred. No. 2.5e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 194 GYVE 198

RESULT 43  
US-10-424-599-173421  
; Sequence 173421, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 173421  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(276)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_127617C.1.pep  
US-10-424-599-173421

Query Match 91.7%; Score 22; DB 15; Length 276;  
Best Local Similarity 80.0%; Pred. No. 2.5e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 69 GYVE 73

RESULT 44  
US-10-369-493-9674  
; Sequence 9674, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 9674



; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Desulfitobacterium hafniense  
US-10-369-493-9674

Query Match 91.7%; Score 22; DB 14; Length 279;  
Best Local Similarity 80.0%; Pred. No. 2.6e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 70 GYSVE 74

## RESULT 45

US-10-136-253-7  
; Sequence 7, Application US/10136253  
; Publication No. US20020136737A1  
; GENERAL INFORMATION:  
; APPLICANT: FRANKEL, Fred R  
; APPLICANT: PORTNOY, Daniel A.  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACIDS COMPRISING LISTERIA DAL AND DAT GENES  
; FILE REFERENCE: 053893-5011-02  
; CURRENT APPLICATION NUMBER: US/10/136,253  
; CURRENT FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: 09/520,207  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: PCT/US98/24357  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: 08/972,902  
; PRIOR FILING DATE: 1997-11-18  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 282  
; TYPE: PRT  
; ORGANISM: Listeria monocytogenes  
US-10-136-253-7

Query Match 91.7%; Score 22; DB 13; Length 282;  
Best Local Similarity 80.0%; Pred. No. 2.6e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 64 GYTVF 68

## RESULT 46

US-10-369-493-23381  
; Sequence 23381, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 23381  
; LENGTH: 296  
; TYPE: PRT  
; ORGANISM: Bacillus subtilis  
US-10-369-493-23381

Query Match 91.7%; Score 22; DB 14; Length 296;

Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 41 GYTVF 45

## RESULT 47

US-10-424-599-283601  
; Sequence 283601, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 283601  
; LENGTH: 296  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_98113C.1.pep  
US-10-424-599-283601

Query Match 91.7%; Score 22; DB 15; Length 296;  
Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 177 GYAVE 181

## RESULT 48

US-10-040-394-9  
; Sequence 9, Application US/10040394  
; Publication No. US20030108641A1  
; GENERAL INFORMATION:  
; APPLICANT: Soe, Jörn  
; APPLICANT: Poulsen, Charlotte  
; APPLICANT: Rasmussen, Preben  
; APPLICANT: Madrig, Susan  
; APPLICANT: Zargahi, Masoud  
; TITLE OF INVENTION: Improved Method for Preparing Flour Doughs and Products Made From  
; TITLE OF INVENTION: Doughs Using a Glycerol Oxidase  
; FILE REFERENCE: 674509-2045.1  
; CURRENT APPLICATION NUMBER: US/10/040,394  
; CURRENT FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: US 09/402,664  
; PRIOR FILING DATE: 1998-04-03  
; PRIOR APPLICATION NUMBER: PCT/DK98/00136  
; PRIOR FILING DATE: 1998-04-03  
; PRIOR APPLICATION NUMBER: DK 0400/97  
; PRIOR FILING DATE: 1997-04-09  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 297  
; TYPE: PRT  
; ORGANISM: Aspergillus tubingensis  
US-10-040-394-9

Query Match 91.7%; Score 22; DB 14; Length 297;  
Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy1 GYXVE 5  
|||

Db190 GYSVE 194

RESULT 49

US-10-150-429B-3

; Sequence 3, Application US/10150429B

; Publication No. US20030175383A1

GENERAL INFORMATION:

APPLICANT: Bojsen, Kirsten

APPLICANT: Poulsen, Charlotte Horsmans

APPLICANT: Soe, Jørn Borch

TITLE OF INVENTION: METHOD OF IMPROVING DOUGH AND BREAD QUALITY

FILE REFERENCE: Y020078

CURRENT APPLICATION NUMBER: US/10/150.429B

CURRENT FILING DATE: 2002-05-17

PRIOR APPLICATION NUMBER: US 60/347,007

PRIOR FILING DATE: 2002-01-09

PRIOR APPLICATION NUMBER: GB 0112226.6

PRIOR FILING DATE: 2001-05-18

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 297

TYPE: PRT

ORGANISM: Aspergillus tubingensis

US-10-150-429B-3

Query Match91.7%; Score 22; DB 14; Length 297;

Best Local Similarity80.0%; Pred. No. 2.7e+03;

Mismatches4; Conservative0; Mismatches1; Indels0; Gaps0;

Qy1 GYXVE 5  
|||

Db190 GYSVE 194

RESULT 50

US-10-462-527-9

; Sequence 9, Application US/10462527

; Publication No. US20040071853A1

GENERAL INFORMATION:

APPLICANT: Soe, Jørn

APPLICANT: Poulsen, Charlotte

APPLICANT: Rasmussen, Preben

APPLICANT: Madrid, Susan

APPLICANT: Zargahi, Masoud

TITLE OF INVENTION: Improved Method for Preparing Flour Doughs and Products Made From

TITLE OF INVENTION: Doughs Using a Glycerol Oxidase

FILE REFERENCE: 674509-2045.2

CURRENT APPLICATION NUMBER: US/10/462,527

CURRENT FILING DATE: 2003-06-16

PRIOR APPLICATION NUMBER: US 10/040,394

PRIOR FILING DATE: 2002-01-09

PRIOR APPLICATION NUMBER: US 09/402,664

PRIOR FILING DATE: 1998-04-03

PRIOR APPLICATION NUMBER: PCT/DK98/00136

PRIOR FILING DATE: 1998-04-03

PRIOR APPLICATION NUMBER: DK 0400/97

PRIOR FILING DATE: 1997-04-09

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn version 3.1

SEQ ID NO 9

LENGTH: 297

TYPE: PRT

ORGANISM: Aspergillus tubingensis

US-10-462-527-9

Query Match91.7%; Score 22; DB 15; Length 297;

Best Local Similarity80.0%; Pred. No. 2.7e+03;

Mismatches4; Conservative0; Mismatches1; Indels0; Gaps0;

Qy1 GYXVE 5  
|||

Db190 GYSVE 194

RESULT 51

US-10-274-694-6

; Sequence 6, Application US/10274694

; Publication No. US20030143589A1

GENERAL INFORMATION:

APPLICANT: BAUGHN, Mariah R.

APPLICANT: BRUNS, Christopher M.

APPLICANT: DAS, Debopriya

APPLICANT: DING, Li

APPLICANT: ELLIOTT, Vicki S.

APPLICANT: GANDHI, Ameena R.

APPLICANT: HAFALIA, April J.A.

APPLICANT: KEARNEY, Liam

APPLICANT: KHAN, Farrah A.

APPLICANT: LAL, Preeti G.

APPLICANT: LEE, Ernestine A.

APPLICANT: LU, Dyung Aina M.

APPLICANT: LU, Yan

APPLICANT: NGUYEN, Dannel B.

APPLICANT: PATTERSON, Chandra S.

APPLICANT: RAMKUMAR, Jayalaxmi

APPLICANT: RING, Huijun Z.

APPLICANT: SANJANWALA, Madhusudan M.

APPLICANT: TANG, Y. Tom

APPLICANT: THANGAVELU, Kavitha

APPLICANT: THORNTON, Michael B.

APPLICANT: TRIBOULEY, Catherine M.

APPLICANT: WALIA, Narinder K.

APPLICANT: XU, Yuming

APPLICANT: YANG, Junming

APPLICANT: YAO, Monique G.

APPLICANT: YUE, Henry

TITLE OF INVENTION: DRUG METABOLIZING ENZYMES

FILE REFERENCE: PI-0151 USA

CURRENT APPLICATION NUMBER: US/10/274,694

CURRENT FILING DATE: 2002-10-18

PRIOR APPLICATION NUMBER: 60/221,837

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: 60/220,037

PRIOR FILING DATE: 2000-07-21

PRIOR APPLICATION NUMBER: 60/218,948

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US01/21324

PRIOR FILING DATE: 2001-07-05

PRIOR APPLICATION NUMBER: 60/216,804

PRIOR FILING DATE: 2000-07-07

NUMBER OF SEQ ID NOS: 36

SOFTWARE: PERL Program

SEQ ID NO 6

LENGTH: 302

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US20030143589A1 7474438CD1

US-10-274-694-6

Query Match91.7%; Score 22; DB 14; Length 302;

Best Local Similarity80.0%; Pred. No. 2.8e+03;

Mismatches4; Conservative0; Mismatches1; Indels0; Gaps0;

Qy1 GYXVE 5  
|||

Db232 GYSVE 236

RESULT 52

US-10-282-122A-71004

```
; Sequence 71004, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71004
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-282-122A-71004

Query Match          91.7%; Score 22; DB 15; Length 307;
Best Local Similarity 80.0%; Pred. No. 2.8e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
Db      256 GYAVE 260

RESULT 53
US-10-282-122A-50846
; Sequence 50846, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
```

```
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50846
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-10-282-122A-50846

Query Match          91.7%; Score 22; DB 15; Length 308;
Best Local Similarity 80.0%; Pred. No. 2.8e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
Db      88 GYSVE 92

RESULT 54
US-10-369-493-122
; Sequence 122, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 122
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Aquifex aeolicus
US-10-369-493-122

Query Match          91.7%; Score 22; DB 14; Length 313;
Best Local Similarity 80.0%; Pred. No. 2.9e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVE 5
Db      111
```

Db 246 GYSVE 250

## RESULT 55

US-09-738-626-3875  
; Sequence 3875, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENO, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 3875  
; LENGTH: 315  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3875

Query Match 91.7%; Score 22; DB 9; Length 315;  
Best Local Similarity 80.0%; Pred. No. 2.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 45 GYSVE 49

## RESULT 56

US-10-152-886-21  
; Sequence 21, Application US/10152886  
; Publication No. US20030064491A1  
; GENERAL INFORMATION:  
; APPLICANT: ECOPIA BIOSCIENCES INC.  
; APPLICANT: Farnet, Chris  
; APPLICANT: Staffa, Alfredo  
; APPLICANT: Zazopoulos, Emmanuel  
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF ENEDIYNE RING  
; FILE REFERENCE: 3011-3US  
; CURRENT APPLICATION NUMBER: US/10/152,886  
; CURRENT FILING DATE: 2002-05-21  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 21  
; LENGTH: 321  
; TYPE: PRT  
; ORGANISM: Micromonospora echinospora callichensis  
US-10-152-886-21

Query Match 91.7%; Score 22; DB 14; Length 321;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||

Db 48 GYAVE 52

## RESULT 57

US-10-282-122A-71334  
; Sequence 71334, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 71334  
; LENGTH: 322  
; TYPE: PRT  
; ORGANISM: Staphylococcus haemolyticus  
US-10-282-122A-71334

Query Match 91.7%; Score 22; DB 15; Length 322;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 271 GYAVE 275

## RESULT 58

US-10-282-122A-71373  
; Sequence 71373, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 71373  
LENGTH: 322  
TYPE: PRT  
ORGANISM: Staphylococcus haemolyticus  
US-10-282-122A-71373

Query Match 91.7%; Score 22; DB 15; Length 322;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 225 GYSVE 229

RESULT 59  
US-10-437-963-159961  
Sequence 159961, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 159961  
LENGTH: 322  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_59286C.1.pep

US-10-437-963-159961

Query Match 91.7%; Score 22; DB 16; Length 322;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 196 GYAVE 200

RESULT 60  
US-10-369-493-229668  
Sequence 229668, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 22968  
LENGTH: 323  
TYPE: PRT  
ORGANISM: Aeropyrum pernix  
US-10-369-493-229668

Query Match 91.7%; Score 22; DB 14; Length 323;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 2 GYAVE 6

RESULT 61  
US-09-815-242-5529  
Sequence 5529, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931

;  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5529  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-5529

Query Match 91.7%; Score 22; DB 9; Length 325;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 228 GYSVE 232

## RESULT 62

US-09-815-242-12354  
; Sequence 12354, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12354  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-12354

Query Match 91.7%; Score 22; DB 9; Length 325;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 228 GYSVE 232

## RESULT 63

US-10-282-122A-43904  
; Sequence 43904, Application US/10282122A  
; Publication No. US20040029129A1

; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 43904  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-282-122A-43904

Query Match 91.7%; Score 22; DB 15; Length 325;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
|||  
Db 228 GYSVE 232

## RESULT 64

US-10-282-122A-70689  
; Sequence 70689, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

RESULT 66  
US-09-815-242-5059  
; Sequence 5059, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 70689  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-282-122A-70689

Query Match 91.7%; Score 22; DB 15; Length 325;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 228 GYSVE 232

RESULT 65  
US-10-369-493-23041  
; Sequence 23041, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 23041  
; LENGTH: 328  
; TYPE: PRT  
; ORGANISM: Bacillus stearothermophilus  
US-10-369-493-23041

Query Match 91.7%; Score 22; DB 14; Length 328;  
Best Local Similarity 80.0%; Pred. No. 3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 175 GYTVE 179

RESULT 67  
US-10-282-122A-43511  
; Sequence 43511, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; QUERY MATCH 91.7%; Score 22; DB 9; Length 344;  
Best Local Similarity 80.0%; Pred. No. 3.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVE 5  
|||  
DB 322 GYAVE 326

; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 43511  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-10-282-122A-43511

Query Match 91.7%; Score 22; DB 15; Length 344;  
Best Local Similarity 80.0%; Pred. No. 3.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 322 GYAVE 326

RESULT 68  
US-10-424-599-272907  
; Sequence 272907, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 272907  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_88458C.1.pgp  
US-10-424-599-272907

Query Match 91.7%; Score 22; DB 15; Length 344;  
Best Local Similarity 80.0%; Pred. No. 3.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 133 GYSVE 137

RESULT 69  
US-10-369-493-21369  
; Sequence 21369, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 21369  
; LENGTH: 345  
; TYPE: PRT  
; ORGANISM: Archaeoglobus fulgidus  
US-10-369-493-21369

Query Match 91.7%; Score 22; DB 14; Length 345;  
Best Local Similarity 80.0%; Pred. No. 3.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 3 GYAVE 7

RESULT 70  
US-10-156-761-12213  
; Sequence 12213, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 12213  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-12213

Query Match 91.7%; Score 22; DB 14; Length 352;  
Best Local Similarity 80.0%; Pred. No. 3.3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5  
Db 105 GYAVE 109

RESULT 71  
US-10-282-122A-62715  
; Sequence 62715, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert



APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 62715

LENGTH: 355

TYPE: PRT

ORGANISM: Mycobacterium bovis

US-10-282-122A-62715

Query Match 91.7%; Score 22; DB 15; Length 355;  
Best Local Similarity 80.0%; Pred. No. 3.3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 13 GYVE 17

#### RESULT 72

US-10-369-493-917

Sequence 917, Application US/10369493

Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 917

LENGTH: 359

TYPE: PRT

ORGANISM: Archaeoglobus fulgidus

US-10-369-493-917

Query Match 91.7%; Score 22; DB 14; Length 359;  
Best Local Similarity 80.0%; Pred. No. 3.3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 292 GYSVE 296

#### RESULT 73

US-10-282-122A-74497

Sequence 74497, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 74497

LENGTH: 360

TYPE: PRT

ORGANISM: Streptococcus pyogenes

US-10-282-122A-74497

Query Match 91.7%; Score 22; DB 15; Length 360;  
Best Local Similarity 80.0%; Pred. No. 3.3e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVE 5

Db 291 GYAVE 295

#### RESULT 74

US-10-282-122A-60356

Sequence 60356, Application US/10282122A

Publication No. US20040029129A1

```
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60356
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60356

Query Match          91.7%; Score 22; DB 15; Length 363;
Best Local Similarity 80.0%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      291 GYAVE 295

RESULT 75
US-10-437-963-147526
; Sequence 147526, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
```

```
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 147526
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48045C.1.pap
US-10-437-963-147526

Query Match          91.7%; Score 22; DB 16; Length 363;
Best Local Similarity 80.0%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVE 5
DB      205 GYAVE 209

Search completed: November 1, 2004, 21:50:49
Job time : 105 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004, CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:14:26 ; Search time 113 Seconds

(without alignments)  
19.048 Million cell updates/sec

Title: US-10-030-194A-5

Perfect score: 28

Sequence: 1 GYXVEE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : A. Genesep23Sep04.\*

1: genesep1980s.\*

2: genesep1990s.\*

3: genesep2000s.\*

4: genesep2001s.\*

5: genesep2002s.\*

6: genesep2003as.\*

7: genesep2003bs.\*

8: genesep2004s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	96.4	45	7	Adf70007 Acma-type
2	27	96.4	235	6	Abm68208 Phototrab
3	27	96.4	282	2	Aay13560 S. haemol
4	27	96.4	344	4	Aau33563 Pseudomon
5	27	96.4	344	6	Abu15587 Protein e
6	27	96.4	349	7	Abu78798 Pseudomon
7	27	96.4	354	7	Adc95552 E. faeciu
8	27	96.4	385	7	Adm25471 Hyperther
9	27	96.4	413	5	Abp26813 Streptoco
10	27	96.4	466	5	Abb54231 Lactococc
11	27	96.4	545	4	Aau34095 Staphyloc
12	27	96.4	553	4	Aau36605 Staphyloc
13	27	96.4	553	6	Abu16137 Protein e
14	27	96.4	553	6	Abm72931 Staphyloc
15	27	96.4	653	8	Adj49360 Oil-assoc
16	27	96.4	762	2	Aay00939 Desulfuro
17	27	96.4	762	5	Aae22116 Desulfuro
18	27	96.4	858	6	Abx53197 Protein s
19	27	96.4	858	7	Adk63242 Disease t
20	27	96.4	1075	5	Abb77436 Human tum
21	27	96.4	1405	6	Abu18105 Protein e
22	27	96.4	1978	2	Aay07032 Breast ca
23	27	96.4	1978	6	Abx64250 Angiogene
24	26	92.9	6	4	Abf31885 Peptide f
25	26	92.9	53	2	Aar43265 RGA1 C-te

26	92.9	54	4	AAQ99655	ERA bindi
27	92.9	64	7	ADM26846	Hyperther
28	92.9	84	4	AAW21351	Peptide #
29	92.9	84	4	ABB43688	Peptide #
30	92.9	84	4	AAW37583	Peptide #
31	92.9	84	4	ABB26636	Protein #
32	92.9	84	4	AAW64647	Human bra
33	92.9	84	4	ABG59061	Human liv
34	92.9	90	4	AAW21427	Peptide #
35	92.9	90	4	ABB43765	Peptide #
36	92.9	90	4	AAW37668	Peptide #
37	92.9	90	4	ABB26708	Protein #
38	92.9	90	4	AAW77498	Human bon
39	92.9	90	4	AAW64727	Human bra
40	92.9	90	4	ABG59136	Human liv
41	92.9	102	2	AAW78901	Human UNC
42	92.9	102	8	ADK16101	Nancarcha
43	92.9	107	7	ADC96202	E. faeciu
44	92.9	112	3	AAW41337	Human ORF
45	92.9	112	5	ABP32756	Human loo
46	92.9	115	3	AAW28420	Corn bran
47	92.9	116	5	ABB49265	Listeria
48	92.9	118	4	AAW48256	Corn magn
49	92.9	128	7	ADH85685	Enterococ
50	92.9	139	4	AAW15152	Peptide #
51	92.9	139	4	AAW27609	Peptide #
52	92.9	139	4	ABB28977	Peptide #
53	92.9	139	4	ABB19587	Protein #
54	92.9	139	4	AAW67316	Human bon
55	92.9	139	4	AAW54935	Human bra
56	92.9	139	4	ABG48978	Human liv
57	92.9	142	8	ABO57215	Human gen
58	92.9	154	3	AAW57021	Human pro
59	92.9	160	4	ABB11168	Human sec
60	92.9	161	5	ABB47842	Listeria
61	92.9	173	8	ABO59796	Human gen
62	92.9	188	4	ABG17723	Novel hum
63	92.9	194	3	AAW43323	Arabidops
64	92.9	197	4	ADG27817	Human nov
65	92.9	199	3	AAW43322	Arabidops
66	92.9	213	4	ABG02504	Novel hum
67	92.9	218	2	AAW13620	Vegetativ
68	92.9	218	2	AAW13619	Vegetativ
69	92.9	218	4	ABG02618	Novel hum
70	92.9	231	8	ADN46899	Thermococ
71	92.9	231	8	ADN46899	Thermococ
72	92.9	233	4	ABB64001	Drosophil
73	92.9	238	7	ADC96641	E. faeciu
74	92.9	239	4	ABG08795	Novel hum
75	92.9	242	5	ABB49186	Listeria
76	92.9	250	4	ABG02628	Novel hum
77	92.9	257	3	AAW43321	Arabidops
78	92.9	257	3	AAW38222	Arabidops
79	92.9	257	3	AAW07071	Arabidops
80	92.9	258	3	AAW04504	Arabidops
81	92.9	260	4	AAW48257	Corn magn
82	92.9	261	4	ABG02502	Novel hum
83	92.9	261	4	ABG02632	Novel hum
84	92.9	270	3	AAW38995	Arabidops
85	92.9	270	3	AAW05652	Arabidops
86	92.9	277	3	AAW07070	Arabidops
87	92.9	277	3	AAW38224	Arabidops
88	92.9	278	3	AAW05651	Arabidops
89	92.9	278	3	AAW38994	Arabidops
90	92.9	286	8	ABO59806	Human gen
91	92.9	297	3	AAW24657	Plant SDF
92	92.9	307	3	AAW28421	Rice bran
93	92.9	318	3	AAW07069	Arabidops
94	92.9	318	3	AAW29240	Arabidops
95	92.9	318	3	AAW38223	Arabidops
96	92.9	319	4	ABG06331	Novel hum
97	92.9	320	3	AAW24903	Plant SDF
98	92.9	321	3	AAW04503	Arabidops

99	26	92.9	323	4	AAU01642	Aau01642 Human sec	172	26	92.9	534	6	ABU42870	Abu42870 Protein e
100	26	92.9	327	6	AAG05650	Arabidops	173	26	92.9	538	4	ABG04019	Novel hum
101	26	92.9	327	3	ABU23487	Protein e	174	26	92.9	539	3	AGS0383	Arabidops
102	26	92.9	331	3	AB24656	Plant SDF	175	26	92.9	540	2	ABY34376	Porphorym
103	26	92.9	332	4	ABG02620	Novel hum	176	26	92.9	550	4	ABB68146	Drosophil
104	26	92.9	333	4	ABG02511	Novel hum	177	26	92.9	554	7	ADM03997	Human pro
105	26	92.9	337	3	AG38993	Arabidops	178	26	92.9	572	4	AB31883	Amino aci
106	26	92.9	343	8	ADQ07572	Streptomy	179	26	92.9	587	3	AB28574	Arabidops
107	26	92.9	345	2	AAW46493	Human ADP	180	26	92.9	587	4	AAE01892	A. thalia
108	26	92.9	351	4	AB27965	Coryneb	181	26	92.9	587	4	AAE02560	Thalecres
109	26	92.9	352	3	AB24655	Plant SDF	182	26	92.9	587	8	ADO01803	Protein e
110	26	92.9	354	3	AB24902	Plant SDF	183	26	92.9	623	7	ADY02540	Rice SLR(
111	26	92.9	354	8	ADN74383	Thale cre	184	26	92.9	623	7	ADC47019	Rice SLR1
112	26	92.9	356	3	AG04502	Arabidops	185	26	92.9	625	8	ADC79243	Rice SLR1
113	26	92.9	357	5	ABP39560	Staphyloc	186	26	92.9	625	8	ADM48144	Polypepti
114	26	92.9	357	7	AD83433	Human Pro	187	26	92.9	630	2	AY02541	Protein e
115	26	92.9	358	4	AD83433	Human Pro	188	26	92.9	630	8	ADM48151	Polypepti
116	26	92.9	358	6	AA88462	S. epider	189	26	92.9	630	8	ADM48151	Human sec
117	26	92.9	358	6	ADA89566	Staphyloc	189	26	92.9	638	3	AY91524	Human pol
118	26	92.9	358	6	ABU43652	Protein e	190	26	92.9	638	4	AAW93431	Human pol
119	26	92.9	358	6	ABU42547	Protein e	191	26	92.9	698	8	ADL71596	Novel hum
120	26	92.9	358	6	ABU42326	Protein e	192	26	92.9	698	8	ADL31029	Human pro
121	26	92.9	358	6	BM72446	Staphyloc	193	26	92.9	710	7	ABO76200	Pseudomon
122	26	92.9	358	7	ADH48621	Experimen	194	26	92.9	712	4	AB79540	Coryneb
123	26	92.9	358	7	ADH48619	Branchied-	195	26	92.9	712	4	AB79540	Malate sy
124	26	92.9	360	3	AG29239	Arabidops	196	26	92.9	739	2	AA77501	Coryneb
125	26	92.9	362	5	ABP40558	Staphyloc	197	26	92.9	739	4	AB79539	Coryneb
126	26	92.9	366	4	AB95186	Human pro	198	26	92.9	752	7	ADB64849	Human pro
127	26	92.9	366	7	ADN95117	Human LEC	199	26	92.9	752	7	ADB64849	Human pro
128	26	92.9	367	4	AB79764	Coryneb	200	26	92.9	791	4	ABG02501	Novel hum
129	26	92.9	367	5	ABB80961	C. glutam	201	26	92.9	791	6	ABG02501	Novel hum
130	26	92.9	367	5	ABB83816	Coryneb	202	26	92.9	791	6	ABG02501	Novel hum
131	26	92.9	370	4	AAU01672	Human sec	203	26	92.9	832	8	ADH22362	Human rec
132	26	92.9	374	3	AAU01672	Human sec	204	26	92.9	832	8	ADH22362	Human rec
133	26	92.9	379	4	AG92160	C. glutam	206	26	92.9	859	7	ADC97202	E. faeciu
134	26	92.9	381	2	ADY1505	C. glutam	207	26	92.9	859	7	ADC97202	E. faeciu
135	26	92.9	381	2	ADY1505	C. glutam	208	26	92.9	865	6	ABU41227	Protein e
136	26	92.9	384	3	AG29238	Arabidops	209	26	92.9	865	6	ABU41227	Protein e
137	26	92.9	384	3	AG29238	Arabidops	210	26	92.9	865	6	ABU41227	Protein e
138	26	92.9	386	4	ABB65853	Drosophil	211	26	92.9	873	8	ADN00738	Human LDL
139	26	92.9	386	4	ABB71579	Drosophil	212	26	92.9	873	8	ADN00738	Human LDL
140	26	92.9	396	3	AA94918	Human sec	213	26	92.9	929	7	ADN00738	Human LDL
141	26	92.9	399	4	AB68540	Human GRP	214	26	92.9	929	7	ADN00738	Human LDL
142	26	92.9	399	5	ABP28499	Streptoco	215	26	92.9	933	5	AAO18735	Human NOV
143	26	92.9	405	3	AG29709	Arabidops	216	26	92.9	933	5	AAO18735	Human NOV
144	26	92.9	412	7	ADP07610	Bacterial	217	26	92.9	945	2	AAW78900	Rat UNC-5
145	26	92.9	413	3	AG29708	Arabidops	218	26	92.9	945	2	AAW78900	Rat UNC-5
146	26	92.9	413	3	AD30985	Plant yie	219	26	92.9	945	5	ABB09520	Human tra
147	26	92.9	413	8	AD14249	Plant tra	220	26	92.9	945	5	ABB09520	Human tra
148	26	92.9	425	2	AA902539	Protein e	221	26	92.9	945	6	ABU80942	Human PRO
149	26	92.9	425	5	ABB97390	Novel hum	222	26	92.9	945	6	ABU80942	Human PRO
150	26	92.9	445	4	ABG28875	Novel hum	223	26	92.9	945	6	ABU66642	Human PRO
151	26	92.9	445	7	ADM26351	Hyperther	224	26	92.9	945	6	ABU66642	Human PRO
152	26	92.9	458	4	AAW25289	Human pro	225	26	92.9	945	6	ABU66642	Human PRO
153	26	92.9	460	5	ABP40477	Staphyloc	226	26	92.9	945	6	ABU66642	Human PRO
154	26	92.9	475	7	AAQ50384	Arabidops	227	26	92.9	945	6	ADAI8746	Human PRO
155	26	92.9	475	7	AAQ50384	Arabidops	228	26	92.9	945	6	ADAI8746	Human PRO
156	26	92.9	477	4	ABG01466	Novel hum	229	26	92.9	945	6	ADAI8746	Human PRO
157	26	92.9	483	4	ABG03737	Novel hum	230	26	92.9	945	6	ADAI8746	Human PRO
158	26	92.9	492	2	AA934503	Porphorym	231	26	92.9	945	6	ADAI8746	Human PRO
159	26	92.9	505	4	ABG27093	Novel hum	232	26	92.9	945	6	ADAI8746	Human PRO
160	26	92.9	511	8	AD061601	Transcrip	233	26	92.9	945	6	ADAI8746	Human PRO
161	26	92.9	517	3	AG38577	Arabidops	234	26	92.9	945	6	ADAI8746	Human PRO
162	26	92.9	518	3	AG38576	Arabidops	235	26	92.9	945	6	ADAI8746	Human PRO
163	26	92.9	531	3	AB28575	Arabidops	236	26	92.9	945	6	ADAI8746	Human PRO
164	26	92.9	532	2	AAW30792	Arabidops	237	26	92.9	945	6	ADAI8746	Human PRO
165	26	92.9	532	4	AAE01907	Arabidops	238	26	92.9	945	6	ADAI8746	Human PRO
166	26	92.9	532	4	AAE02545	A. thalia	239	26	92.9	945	6	ADAI8746	Human PRO
167	26	92.9	532	7	ADD55688	Thalecres	240	26	92.9	945	6	ADAI8746	Human PRO
168	26	92.9	532	7	ADD55688	Thalecres	241	26	92.9	945	6	ADAI8746	Human PRO
169	26	92.9	532	8	AD143827	Plant tra	242	26	92.9	945	6	ADAI8746	Human PRO
170	26	92.9	532	8	AD001805	Thalecres	243	26	92.9	945	6	ADAI8746	Human PRO
171	26	92.9	533	3	AG38575	Arabidops	244	26	92.9	945	6	ADAI8746	Human PRO

245	26	92.9	945	6	ADB19706	ADB19706 Novel hum	318	26	92.9	945	7	ADD09800	ADD09800 Human PRO
246	26	92.9	945	6	ADB13018	ADB13018 Human PRO	319	26	92.9	945	7	ADD04375	ADD04375 Novel hum
247	26	92.9	945	6	ABO43221	Novel hum	320	26	92.9	945	7	ADC80331	ADC80331 Novel hum
248	26	92.9	945	6	ADA74272	Human PRO	321	26	92.9	945	7	ADD10838	ADD10838 Human PRO
249	26	92.9	945	6	ADB24505	Human PRO	322	26	92.9	945	7	ADC47719	ADC47719 Human PRO
250	26	92.9	945	6	ADA82029	Human PRO	323	26	92.9	945	7	ADC79779	ADC79779 Novel hum
251	26	92.9	945	6	ADA74992	Human PRO	324	26	92.9	945	7	ADD09248	ADD09248 Human PRO
252	26	92.9	945	6	ADA85070	Novel hum	325	26	92.9	945	7	ADD40961	ADD40961 Novel hum
253	26	92.9	945	6	ADA84518	Novel hum	326	26	92.9	945	7	ADD52100	ADD52100 Human PRO
254	26	92.9	945	6	ADB29774	Human PRO	327	26	92.9	945	7	ADD52840	ADD52840 Human PRO
255	26	92.9	945	6	ADA80302	Human PRO	328	26	92.9	945	7	ADD53392	ADD53392 Novel hum
256	26	92.9	945	6	ADA75544	Human PRO	329	26	92.9	945	7	ADD51548	ADD51548 Human PRO
257	26	92.9	945	6	ADA46769	Human PRO	330	26	92.9	945	7	ADD02347	ADD02347 Human PRO
258	26	92.9	945	6	ADB25065	Human PRO	331	26	92.9	945	7	ADD01781	ADD01781 Human PRO
259	26	92.9	945	6	ADA93241	Human PRO	332	26	92.9	945	7	ADD53963	ADD53963 Novel hum
260	26	92.9	945	6	ADB26591	Human PRO	333	26	92.9	945	7	ADD63096	ADD63096 Rat Prote
261	26	92.9	945	6	ADB30878	Human PRO	334	26	92.9	945	7	ADD92280	ADD92280 Human PRO
262	26	92.9	945	6	ADA60806	Homo sapi	335	26	92.9	945	7	ADD91176	ADD91176 Human PRO
263	26	92.9	945	6	ADB23953	Human PRO	336	26	92.9	945	7	ADE03790	ADE03790 Human PRO
264	26	92.9	945	6	ADA96282	Human PRO	337	26	92.9	945	7	ADE32087	ADE32087 Novel hum
265	26	92.9	945	6	ADA80854	Human PRO	338	26	92.9	945	7	ADE22019	ADE22019 Human PRO
266	26	92.9	945	6	ADA95730	Human PRO	339	26	92.9	945	7	ADD79243	ADD79243 Human PRO
267	26	92.9	945	6	ADB26039	Human PRO	340	26	92.9	945	7	ADE41779	ADE41779 Human PRO
268	26	92.9	945	6	ADB21524	Novel hum	341	26	92.9	945	7	ADE17596	ADE17596 Human PRO
269	26	92.9	945	7	ADA77303	Human PRO	342	26	92.9	945	7	ADD91728	ADD91728 Human PRO
270	26	92.9	945	7	ADB18043	Human PRO	343	26	92.9	945	7	ADE33191	ADE33191 Novel hum
271	26	92.9	945	7	ADA86726	Novel hum	344	26	92.9	945	7	ADE33743	ADE33743 Novel hum
272	26	92.9	945	7	ADA87829	Novel hum	345	26	92.9	945	7	ADD97955	ADD97955 Human PRO
273	26	92.9	945	7	ADA46217	Novel hum	346	26	92.9	945	7	ADD92832	ADD92832 Human PRO
274	26	92.9	945	7	ADB28247	Human PRO	347	26	92.9	945	7	ADE19252	ADE19252 Human PRO
275	26	92.9	945	7	ADB28799	Human PRO	348	26	92.9	945	7	ADE18700	ADE18700 Human PRO
276	26	92.9	945	7	ADA76751	Human PRO	349	26	92.9	945	7	ADE42896	ADE42896 Human PRO
277	26	92.9	945	7	ADA88381	Novel hum	350	26	92.9	945	7	ADD95685	ADD95685 Human PRO
278	26	92.9	945	7	ADA97386	Human PRO	351	26	92.9	945	7	ADE22571	ADE22571 Human PRO
279	26	92.9	945	7	ADB27143	Human PRO	352	26	92.9	945	7	ADD78689	ADD78689 Human PRO
280	26	92.9	945	7	ADB22076	Novel hum	353	26	92.9	945	7	ADE32639	ADE32639 Novel hum
281	26	92.9	945	7	ADA66767	Human PRO	354	26	92.9	945	7	ADE42331	ADE42331 Human PRO
282	26	92.9	945	7	ADB22628	Human PRO	355	26	92.9	945	7	ADD80347	ADD80347 Human PRO
283	26	92.9	945	7	ADB23401	Human PRO	356	26	92.9	945	7	ADD89375	ADD89375 Human PRO
284	26	92.9	945	7	ADA92123	Novel hum	357	26	92.9	945	7	ADE40659	ADE40659 Human PRO
285	26	92.9	945	7	ADB15186	Human PRO	358	26	92.9	945	7	ADE04458	ADE04458 Human PRO
286	26	92.9	945	7	ADB38438	Novel hum	359	26	92.9	945	7	ADE92587	ADE92587 Human PRO
287	26	92.9	945	7	ADB37886	Novel hum	360	26	92.9	945	7	ADG21296	ADG21296 Novel hum
288	26	92.9	945	7	ADB66358	Novel hum	361	26	92.9	945	7	ADG22937	ADG22937 Novel hum
289	26	92.9	945	7	ADB89438	Human PRO	362	26	92.9	945	7	ADF97272	ADF97272 Human PRO
290	26	92.9	945	7	ADB90170	Human PRO	363	26	92.9	945	7	ADG80336	ADG80336 Human PRO
291	26	92.9	945	7	ADB39271	Novel hum	364	26	92.9	945	7	ADG79784	ADG79784 Human PRO
292	26	92.9	945	7	ADB46894	Novel hum	365	26	92.9	945	7	ADH55076	ADH55076 Novel hum
293	26	92.9	945	7	ADB86501	Human PRO	366	26	92.9	945	7	ADH55628	ADH55628 Novel hum
294	26	92.9	945	7	ADB77106	Novel hum	367	26	92.9	945	7	ADI64796	ADI64796 Novel hum
295	26	92.9	945	7	ADB34263	Human PRO	368	26	92.9	945	7	ADI63295	ADI63295 Novel hum
296	26	92.9	945	7	ADB35367	Human PRO	369	26	92.9	945	7	ADH81709	ADH81709 Novel hum
297	26	92.9	945	7	ADB33711	Human PRO	370	26	92.9	945	7	ADH81157	ADH81157 Novel hum
298	26	92.9	945	7	ADB34815	Human PRO	371	26	92.9	945	7	ADM82326	ADM82326 Novel hum
299	26	92.9	945	7	ADB35919	Human PRO	372	26	92.9	945	7	ADN15725	ADN15725 Novel hum
300	26	92.9	945	7	ADB46314	Novel hum	373	26	92.9	945	7	ADN16354	ADN16354 Novel hum
301	26	92.9	945	7	ADC50187	Novel hum	374	26	92.9	945	7	ADN15173	ADN15173 Novel hum
302	26	92.9	945	7	ADC71734	Novel hum	375	26	92.9	945	7	ADN14621	ADN14621 Novel hum
303	26	92.9	945	7	ADC59713	Novel hum	376	26	92.9	945	8	ADC80883	ADC80883 Novel hum
304	26	92.9	945	7	ADC52720	Novel hum	377	26	92.9	945	8	ADD76331	ADD76331 Human PRO
305	26	92.9	945	7	ADC57074	Novel hum	378	26	92.9	945	8	ADD87695	ADD87695 Human PRO
306	26	92.9	945	7	ADC60265	Novel hum	379	26	92.9	945	8	ADD86099	ADD86099 Human PRO
307	26	92.9	945	7	ADC50740	Novel hum	380	26	92.9	945	8	ADE75547	ADE75547 Human PRO
308	26	92.9	945	7	ADC65267	Human PRO	381	26	92.9	945	8	ADE23123	ADE23123 Human PRO
309	26	92.9	945	7	ADC54365	Novel hum	382	26	92.9	945	8	ADE23675	ADE23675 Human PRO
310	26	92.9	945	7	ADC53326	Novel hum	383	26	92.9	945	8	ADE24318	ADE24318 Human PRO
311	26	92.9	945	7	ADC58849	Novel hum	384	26	92.9	945	8	ADD87143	ADD87143 Human PRO
312	26	92.9	945	7	ADC55727	Novel hum	385	26	92.9	945	8	ADE89009	ADE89009 Human PRO
313	26	92.9	945	7	ADC58297	Novel hum	386	26	92.9	945	8	ADE18148	ADE18148 Human PRO
314	26	92.9	945	7	ADD02971	Novel hum	387	26	92.9	945	8	ADE88457	ADE88457 Human PRO
315	26	92.9	945	7	ADC89963	Novel hum	388	26	92.9	945	8	ADE94477	ADE94477 Human PRO
316	26	92.9	945	7	ADC69382	Human PRO	389	26	92.9	945	8	ADE90888	ADE90888 Human PRO
317	26	92.9	945	7	ADC48271	Human PRO	390	26	92.9	945	8	ADE95029	ADE95029 Human PRO

391	26	92.9	945	8	ADE93139	Human PRO	464	26	92.9	1104	3	AAG50662	Arabidops
392	26	92.9	945	8	ADF34720	Human PRO	465	26	92.9	1148	3	AGS50661	Arabidops
393	26	92.9	945	8	ADE92035	Novel hum	466	26	92.9	1234	8	ADJ35090	Xylanase
394	26	92.9	945	8	ADE90336	Human PRO	467	26	92.9	1381	8	ADM57193	A thalian
395	26	92.9	945	8	ADE91483	Novel hum	468	26	92.9	1690	3	AAB42248	Human ORF
396	26	92.9	945	8	ADG02062	Human PRO	469	26	92.9	1749	4	AAM79130	Human pro
397	26	92.9	945	8	ADG21848	Novel hum	470	25	89.3	82	7	ADF30922	Soil meta
398	26	92.9	945	8	ADG19918	Human PRO	471	25	89.3	91	4	AUA48276	Propionib
399	26	92.9	945	8	ADF97824	Human PRO	472	25	89.3	91	6	ABM44795	Propionib
400	26	92.9	945	8	ADG24041	Novel hum	473	25	89.3	109	2	AAM20117	H. pylori
401	26	92.9	945	8	ADF98395	Human PRO	474	25	89.3	144	3	AGG09789	Arabidops
402	26	92.9	945	8	ADG03226	Human PRO	475	25	89.3	144	3	AGG10643	Arabidops
403	26	92.9	945	8	ADF98947	Human PRO	476	25	89.3	174	7	ADC96352	E. fasciu
404	26	92.9	945	8	ADG16532	Human PRO	477	25	89.3	178	7	ADH85524	Enterococ
405	26	92.9	945	8	ADG04991	Human PRO	478	25	89.3	179	5	ABUS2054	Helicobac
406	26	92.9	945	8	ADG19258	Human PRO	479	25	89.3	180	2	AAM20190	H. pylori
407	26	92.9	945	8	ADG13095	Human PRO	480	25	89.3	180	6	ABU31138	Protein e
408	26	92.9	945	8	ADG08152	Novel hum	481	25	89.3	185	5	ABUS2002	Helicobac
409	26	92.9	945	8	ADG15322	Human PRO	482	25	89.3	189	2	ABU20690	H. pylori
410	26	92.9	945	8	ADF96720	Human PRO	483	25	89.3	199	6	ABU25830	Protein e
411	26	92.9	945	8	ADG05905	Human PRO	484	25	89.3	202	5	ABB54543	Lactococc
412	26	92.9	945	8	ADG23489	Novel hum	485	25	89.3	225	6	ABU43526	Protein e
413	26	92.9	945	8	ADG03778	Human PRO	486	25	89.3	234	8	ADK16089	Nanoarcha
414	26	92.9	945	8	ADG24679	Novel hum	487	25	89.3	240	8	ADK46964	Streptoco
415	26	92.9	945	8	ADG06976	Novel hum	488	25	89.3	242	3	AGG10642	Arabidops
416	26	92.9	945	8	ADG07528	Novel hum	489	25	89.3	242	3	AGG09788	Arabidops
417	26	92.9	945	8	ADG55023	Novel hum	490	25	89.3	244	3	AY81721	Streptoco
418	26	92.9	945	8	ADG06087	Novel hum	491	25	89.3	244	6	ABU01873	S. pneumo
419	26	92.9	945	8	ADG61791	Novel hum	492	25	89.3	246	3	AGG09787	Arabidops
420	26	92.9	945	8	ADG81992	Human PRO	493	25	89.3	246	3	AGG10641	Arabidops
421	26	92.9	945	8	ADG57231	Novel hum	494	25	89.3	249	4	AAB96686	Putative
422	26	92.9	945	8	ADG56679	Novel hum	495	25	89.3	249	4	ADN47546	Thermococ
423	26	92.9	945	8	ADG55575	Novel hum	496	25	89.3	271	6	ABU43886	Thermococ
424	26	92.9	945	8	ADG58335	Novel hum	497	25	89.3	286	6	ABR53057	Protein s
425	26	92.9	945	8	ADG70701	Novel hum	498	25	89.3	286	7	ADK62930	Disease t
426	26	92.9	945	8	ADG57783	Novel hum	499	25	89.3	300	4	AAB96658	Putative
427	26	92.9	945	8	ADG53367	Novel hum	500	25	89.3	300	8	ADN46691	Thermococ
428	26	92.9	945	8	ADG71253	Novel hum	501	25	89.3	306	8	ADN46838	Thermococ
429	26	92.9	945	8	ADG81440	Human PRO	502	25	89.3	310	3	ADE80899	Escherich
430	26	92.9	945	8	ADH30402	Human PRO	503	25	89.3	331	4	AAU56366	Propionib
431	26	92.9	945	8	ADH11769	Novel hum	504	25	89.3	331	5	ABBA48211	Listeria
432	26	92.9	945	8	ADG52191	Novel hum	505	25	89.3	331	5	ABM52885	Propionib
433	26	92.9	945	8	ADG53919	Novel hum	506	25	89.3	335	2	AAW21007	H. pylori
434	26	92.9	945	8	ADG80888	Human PRO	507	25	89.3	335	7	ABO77672	Pseudomon
435	26	92.9	945	8	ADG56127	Novel hum	508	25	89.3	349	8	ADP30135	Human sec
436	26	92.9	945	8	ADH12393	Novel hum	509	25	89.3	352	7	ADM25397	Hyperther
437	26	92.9	945	8	ADG61239	Novel hum	510	25	89.3	353	8	ADP30134	Human sec
438	26	92.9	945	8	ADH28326	Human PRO	511	25	89.3	375	6	ABU11647	Human MDD
439	26	92.9	945	8	ADG54471	Novel hum	512	25	89.3	398	4	ABB63011	Drosophil
440	26	92.9	945	8	ADG59511	Novel hum	513	25	89.3	432	6	ABU28600	Protein e
441	26	92.9	945	8	ADH80935	Human PRO	514	25	89.3	440	6	ADN47614	Thermococ
442	26	92.9	945	8	ADG09678	Novel hum	515	25	89.3	442	6	ABU24620	Protein e
443	26	92.9	945	8	ADH15149	Novel hum	516	25	89.3	480	5	ABB92725	Herbicida
444	26	92.9	945	8	ADG09026	Novel hum	517	25	89.3	514	5	ABP27404	Streptoco
445	26	92.9	945	8	ADH14481	Novel hum	518	25	89.3	522	7	ABO63524	Klebsiell
446	26	92.9	945	8	ADH18076	Novel hum	519	25	89.3	546	4	ABB65311	Drosophil
447	26	92.9	945	8	ADJ63357	Novel hum	520	25	89.3	678	6	ABG29662	Novel hum
448	26	92.9	945	8	ADJ77252	Human PRO	521	25	89.3	681	6	ABU45148	Protein e
449	26	92.9	945	8	ADJ65374	Human PRO	522	25	89.3	865	4	AAU34558	E. coli c
450	26	92.9	945	8	ADM27510	Human PRO	523	25	89.3	865	6	ABU28615	Protein e
451	26	92.9	945	8	ADM42234	Human PRO	524	25	89.3	865	6	ABU31554	Protein e
452	26	92.9	945	8	ADM28096	Human PRO	525	25	89.3	865	6	ABU338255	Protein e
453	26	92.9	946	5	ABG61795	Novel UNC	526	25	89.3	867	4	AAU338255	Protein e
454	26	92.9	950	7	ADN40022	Cancer/an	527	25	89.3	868	4	AAU35640	Haemophil
455	26	92.9	1024	5	ABP26425	Streptoco	528	25	89.3	868	6	AAU33626	Pseudomon
456	26	92.9	1037	3	AGS50663	Arabidops	529	25	89.3	868	6	ABU15524	Protein e
457	26	92.9	1058	6	ABU46574	Protein e	530	25	89.3	868	6	ABU15524	Protein e
458	26	92.9	1059	6	ABU44367	Protein e	531	25	89.3	869	6	ABU39578	Protein e
459	26	92.9	1060	6	ABU29151	Protein e	532	25	89.3	869	6	ABU41607	Protein e
460	26	92.9	1061	6	ABU29728	Protein e	533	25	89.3	869	7	ABO61210	Klebsiell
461	26	92.9	1074	7	ADH88205	Enterococ	534	25	89.3	889	7	ABO69956	Pseudomon
462	26	92.9	1088	7	ADJ70163	Human hea	535	25	89.3	912	2	AAW42103	Human p11
463	26	92.9	1094	8	ADJ34928	Xylanase	536	25	89.3	912	2	AAW41009	Amino aci

537	25	89.3	919	2	AA41011	Aay41011 Amino aci	610	24	85.7	381	4	AA879552	Aab79552 Corynebac
538	25	89.3	966	4	AB869689	Abd56889 Drosophil	611	24	85.7	383	4	AA990479	Aag90479 C glutami
539	25	89.3	1086	7	AD59735	Rat Prote	612	24	85.7	394	4	AA990519	Aag90519 C glutami
540	25	89.3	1086	7	AD45680	Abd45680 Rat Prote	613	24	85.7	396	6	AB073753	Abu43753 Protein e
541	25	89.3	2000	6	ABR53747	Abt53747 Protein s	614	24	85.7	406	6	ABD07130	Abd07130 Alloioococ
542	25	89.3	2000	7	ADK64668	Adk64668 Disease t	615	24	85.7	407	6	ABD08098	Abd08098 Alloioococ
543	25	89.3	2951	4	AB860291	Abb60291 Drosophil	616	24	85.7	424	8	ADH12917	Adh12917 Francisel
544	24	85.7	13	2	AA42420	Aar42420 Amino aci	617	24	85.7	449	6	ADU20343	Adu20343 Protein e
545	24	85.7	13	2	AAW74709	Aaw74709 Peptide 3	618	24	85.7	503	8	ADM48123	Adm48123 Polyypepti
546	24	85.7	13	3	AA59530	Aay59530 Taste mod	619	24	85.7	509	5	ABB92085	Abb92085 Herbicida
547	24	85.7	13	6	ABU61589	Abu61589 Bovine ro	620	24	85.7	513	4	AAU35015	Aau35015 Enterococ
548	24	85.7	46	2	AAW75574	Aaw75574 M. tuberc	621	24	85.7	514	5	ABB92726	Abb92726 Herbicida
549	24	85.7	46	2	AAW75573	Aaw75573 M. tuberc	622	24	85.7	514	8	ADQ60096	Adq60096 Arabidops
550	24	85.7	47	2	AA92887	Aar92887 Mycobacte	623	24	85.7	523	7	ADH88121	Adh88121 Enterococ
551	24	85.7	47	6	ABG74428	Abg74428 M. tuberc	624	24	85.7	528	7	ADC96713	Adc96713 E. faeciu
552	24	85.7	60	3	AA601908	Aag01908 Human sec	625	24	85.7	532	7	ADF07329	Adf07329 Bacterial
553	24	85.7	64	7	ADC31737	Aac31737 Human nov	626	24	85.7	552	6	ABU17992	Abu17992 Protein e
554	24	85.7	74	8	ADI10310	Adi10310 Protein e	627	24	85.7	553	7	ADC00328	Adc00328 Enterohae
555	24	85.7	96	7	ADH88231	Adh88231 Enterococ	628	24	85.7	553	7	ADC00856	Adc00856 Enterohae
556	24	85.7	101	4	AAO08363	Aao08363 Human pol	629	24	85.7	553	7	ADC00054	Adc00054 Enterohae
557	24	85.7	101	5	ABP39018	Abp39018 Staphyloc	630	24	85.7	553	7	ADC00481	Adc00481 Enterohae
558	24	85.7	111	4	ABG28470	Abg28470 Novel hum	631	24	85.7	576	6	ABM65451	Abm65451 Propionib
559	24	85.7	116	7	ADC33292	Adc33292 Human nov	632	24	85.7	592	5	ABP73559	Abp73559 Candida a
560	24	85.7	133	7	ADC87832	Adc87832 Ribosomal	633	24	85.7	627	2	AAW85597	Aaw85597 Human GAB
561	24	85.7	136	7	ADC89252	Adc89252 Ribosomal	634	24	85.7	627	2	AAW85596	Aaw85596 Human GAB
562	24	85.7	143	4	AA882310	Aag82310 S. epider	635	24	85.7	627	6	ABJ19813	Abj19813 Androgen-
563	24	85.7	143	6	ABU42792	Abu42792 Protein e	636	24	85.7	630	4	AAU48147	Aau48147 Propionib
564	24	85.7	143	6	ABU43714	Abu43714 Protein e	637	24	85.7	630	6	ABM44666	Abm44666 Propionib
565	24	85.7	150	4	AB868217	Abb68217 Drosophil	638	24	85.7	726	6	ABU24278	Abu24278 Protein e
566	24	85.7	152	7	ADC89251	Adc89251 Ribosomal	639	24	85.7	853	5	ABP73993	Abp73993 Candida a
567	24	85.7	194	7	ABO74148	Abu74148 Pseudomon	640	24	85.7	865	5	ABG30569	Abg30569 Glycine m
568	24	85.7	198	3	AA808773	Aag08773 Arabidops	641	24	85.7	896	4	ABG28816	Abg28816 Novel hum
569	24	85.7	198	3	AA646166	Aag46166 Arabidops	642	24	85.7	1040	6	ABU20973	Abu20973 Protein e
570	24	85.7	200	7	ADH86317	Adh86317 Enterococ	643	24	85.7	1046	6	ABU31607	Abu31607 Protein e
571	24	85.7	201	4	AA931162	Aag931162 C glutami	644	24	85.7	1091	5	ABB48258	Abb48258>Listeria
572	24	85.7	206	4	AA879546	Aab79546 Corynebac	645	24	85.7	1095	7	ABO62750	Abu62750 Klebsiell
573	24	85.7	217	5	AB854592	Abb54592 Lactococc	646	24	85.7	1572	4	AA8682906	Aag82906 S. epider
574	24	85.7	224	5	ABP28289	Abp28289 Streptoco	647	24	85.7	1716	7	ADF07324	Adf07324 Bacterial
575	24	85.7	224	5	AAO14749	Aao14749 Corynebac	648	24	85.7	1972	5	AA878488	Aag78488 Human p53
576	24	85.7	240	7	ABM62674	Adm26274 Hyperther	649	24	85.7	1972	8	ADO30967	Ado30967 Human Pol
577	24	85.7	253	6	ABM68058	Abm68058 Photorhab	650	24	85.7	2618	4	ABG02135	Abg02135 Novel hum
578	24	85.7	263	3	AA808772	Aag08772 Arabidops	651	24	85.7	2622	4	ABG06418	Abg06418 Novel hum
579	24	85.7	263	3	AA646165	Aag46165 Arabidops	652	23	82.1	9	2	AAW87934	Aaw87934 Lucilia c
580	24	85.7	264	4	AA81841	Aag81841 S. epider	653	23	82.1	10.	8	ADK05725	Adk05725 Hepatitis
581	24	85.7	264	5	ABP65018	Abp65018 Human pro	654	23	82.1	10	8	ADK05726	Adk05726 Hepatitis
582	24	85.7	268	5	ABP39884	Abp39884 Staphyloc	655	23	82.1	18	6	ABP82141	Abp82141 G protein
583	24	85.7	284	7	ABO84074	Abu84074 Pseudomon	656	23	82.1	23	5	ABG69123	Abg69123 M-2 ligand
584	24	85.7	295	4	AB860698	Abb60698 Drosophil	657	23	82.1	32	3	AA857603	AA857603 Arabidops
585	24	85.7	296	7	ADF41714	Adf41714 Bacillus	658	23	82.1	38	8	ADK01884	Adk01884 Hepatitis
586	24	85.7	310	3	AA836244	Aab36244 T. pallid	659	23	82.1	40	2	AAW27908	Aaw27908 Staphyloc
587	24	85.7	313	7	ADM26457	Adm26457 Hyperther	660	23	82.1	51	4	AAU47305	Aau47305 Propionib
588	24	85.7	315	4	AA873946	Aab73946 Fusion pr	661	23	82.1	51	6	ABM43824	Abm43824 Propionib
589	24	85.7	334	7	ABO62811	Abu62811 Klebsiell	662	23	82.1	58	3	AA844956	AA844956 Zea mays
590	24	85.7	345	3	AA810348	Aag10348 Arabidops	663	23	82.1	68	6	ABR81756	ABr81756 MPV1 V O
591	24	85.7	347	4	AAU58802	Aau58802 Propionib	664	23	82.1	70	7	ADC95659	Adc95659 E. faeciu
592	24	85.7	347	6	ABM55321	Abm55321 Propionib	665	23	82.1	75	4	ABB15710	Abb15710 Human ner
593	24	85.7	350	2	AA42426	Aar42426 Bovine ro	666	23	82.1	76	5	ABP09594	Abp09594 Human ORF
594	24	85.7	350	3	AA595336	Aay595336 Bovine ro	667	23	82.1	77	2	AA31459	Aay31459 A. thalia
595	24	85.7	350	4	AB899067	Ab899067 Human G-p	668	23	82.1	77	3	AA25818	AA25818 ANT-R1 di
596	24	85.7	350	5	AB809278	Ab809278 G protein	669	23	82.1	77	4	AAW90614	Aaw90614 Human imm
597	24	85.7	350	5	AB877787	Ab877787 Amino aci	670	23	82.1	78	7	ADG88449	Adg88449 Arabidops
598	24	85.7	350	6	ABU61595	Abu61595 Rat rod t	671	23	82.1	81	5	ABP01866	Abp01866 Human ORF
599	24	85.7	350	6	ABR82633	ABr82633 C. elegan	672	23	82.1	84	5	ABP03285	Abp03285 Human ORF
600	24	85.7	350	7	ADC09613	Adc09613 Human G-p	673	23	82.1	86	4	ABG07436	Abg07436 Novel hum
601	24	85.7	350	7	ADJ68622	Adj68622 Human hea	674	23	82.1	86	5	ABG69477	ABg69477 Human bai
602	24	85.7	350	8	ADN06142	Adn06142 Bovine Gt	675	23	82.1	87	4	AAU58059	Aau58059 Propionib
603	24	85.7	351	6	ABU23669	Abu23669 Protein e	676	23	82.1	87	6	ABM54578	Abm54578 Propionib
604	24	85.7	351	7	ADC95066	Adc95066 E. faeciu	677	23	82.1	92	5	ABU67208	Abu67208 G-protein
605	24	85.7	352	7	ADM26449	Adm26449 Hyperther	678	23	82.1	92	6	ABP53989	Abp53989 Human bet
606	24	85.7	358	6	ABU29877	Abu29877 Protein e	679	23	82.1	92	8	ADO28729	Ado28729 Human bet
607	24	85.7	358	6	ABU29150	Abu29150 Protein e	680	23	82.1	92	8	ADO05233	Ado05233 Beta-2-ad
608	24	85.7	371	7	ADH88204	Adh88204 Enterococ	681	23	82.1	96	5	AB853948	AB853948 Lactococc
609	24	85.7	374	6	ABM64688	Abm64688 Propionib	682	23	82.1	96	5	ABB53444	ABB53444 Lactococc

683	23	82.1	96	5	ABB54139	Abb54139 Lactococc	756	23	82.1	206	3	AAO7695	Aag07695 Arabidops
684	23	82.1	96	5	ABB55428	Abb55428 Lactococc	757	23	82.1	207	6	ABU26119	Abu26119 Protein e
685	23	82.1	96	5	ABB554136	Abb554136 Lactococc	758	23	82.1	208	3	AG39024	Aag39024 Arabidops
686	23	82.1	96	5	ABB53676	Abb53676 Lactococc	759	23	82.1	209	3	AG36752	Aag36752 Arabidops
687	23	82.1	96	5	ABB53439	Abb53439 Lactococc	760	23	82.1	209	7	ADD30372	Add30372 Plant Vie
688	23	82.1	96	5	ABB55500	Abb55500 Lactococc	761	23	82.1	209	8	ADI44299	Adi44299 Plant tra
689	23	82.1	96	5	ABB53354	Abb53354 Lactococc	762	23	82.1	211	4	AB19434	Abi19434 Nucleotid
690	23	82.1	99	3	AG34449	Aag34449 Arabidops	763	23	82.1	211	4	AG65508	Aag65508 Amino aci
691	23	82.1	103	3	AG54132	Aag54132 Arabidops	764	23	82.1	211	6	AG93184	Abu93184 Tetracycl
692	23	82.1	103	3	AG26980	Aag26980 Arabidops	765	23	82.1	211	7	ADF38930	Adf38930 Transposo
693	23	82.1	103	7	ADC88510	Adc88510 Ribosomal	766	23	82.1	212	6	AG68880	Abm68880 Photorhab
694	23	82.1	105	6	ABU36512	Abu36512 Protein e	767	23	82.1	212	6	ABU40371	Abu40371 Protein e
695	23	82.1	105	6	ABU36025	Abu36025 Protein e	768	23	82.1	214	2	AAW61022	AAW61022 Streptoco
696	23	82.1	105	6	ABU34570	Abu34570 Protein e	769	23	82.1	214	5	ABP30501	ABP30501 Streptoco
697	23	82.1	105	6	ADC88524	Adc88524 Ribosomal	770	23	82.1	214	5	ADB11412	ADB11412 Alloiococ
698	23	82.1	108	4	ABG07141	Abg07141 Novel hum	771	23	82.1	215	7	ADF38957	Adf38957 Agrobacte
699	23	82.1	110	7	ADG88447	Adg88447 Arabidops	772	23	82.1	215	2	AY08492	AY08492 Rat serin
700	23	82.1	111	4	ABG07149	Abg07149 Novel hum	773	23	82.1	216	3	AG08402	Aag08402 Arabidops
701	23	82.1	114	3	AG44954	Aag44954 Zea mays	774	23	82.1	218	2	AY13622	AY13622 Vegetativ
702	23	82.1	115	4	ABG07133	Abg07133 Novel hum	775	23	82.1	218	2	AY13621	AY13621 Vegetativ
703	23	82.1	123	3	AG19181	Aag19181 Arabidops	776	23	82.1	221	4	ABG07130	Abg07130 Novel hum
704	23	82.1	124	4	AAO06786	AAO06786 Human pol	777	23	82.1	221	6	ABP78332	Abp78332 N. gonorr
705	23	82.1	124	8	ADP09946	ADP09946 Anti-mous	778	23	82.1	221	8	ADP08249	Adp08249 Neisseria
706	23	82.1	124	8	ADO25696	ADO25696 Anti-TNF-	779	23	82.1	223	6	ABP78332	Abp78332 N. gonorr
707	23	82.1	125	4	ABG26639	Abg26639 Novel hum	780	23	82.1	225	5	ABG07130	Abg07130 Novel hum
708	23	82.1	125	7	ABO79856	ABO79856 Pseudomon	781	23	82.1	225	5	ABG07130	Abg07130 Novel hum
709	23	82.1	130	3	AAV94439	AAV94439 Soybean N	782	23	82.1	225	3	AG36751	Aag36751 Arabidops
710	23	82.1	133	4	ABG04942	Abg04942 Novel hum	783	23	82.1	226	3	AG36751	Aag36751 Arabidops
711	23	82.1	136	5	ADK36821	Adk36821 Novel hum	784	23	82.1	226	4	ABG04513	Abg04513 Novel hum
712	23	82.1	138	3	AG54131	Aag54131 Arabidops	785	23	82.1	226	5	ABP30425	Abp30425 Streptoco
713	23	82.1	139	4	AA883975	AA883975 S. epider	786	23	82.1	228	4	ABG17232	Abg17232 Novel hum
714	23	82.1	139	4	ABG83104	ABG83104 S. epider	787	23	82.1	228	5	ABP27179	ABP27179 Streptoco
715	23	82.1	139	4	ABG26286	Abg26286 Novel hum	788	23	82.1	228	5	ABP26598	ABP26598 Streptoco
716	23	82.1	139	4	ABG07147	Abg07147 Novel hum	789	23	82.1	228	5	ABP39528	ABP39528 Staphyloc
717	23	82.1	140	3	AG45402	Aag45402 Arabidops	790	23	82.1	230	7	ADL18517	Adl18517 Rice glut
718	23	82.1	140	5	ABG69478	ABG69478 Human bai	791	23	82.1	231	7	ADL18517	Adl18517 Rice glut
719	23	82.1	148	4	ABG07131	Abg07131 Novel hum	792	23	82.1	232	4	ABG24580	Abg24580 Novel hum
720	23	82.1	153	4	AAW90901	AAW90901 Human imm	793	23	82.1	232	5	ABP25449	ABP25449 Streptoco
721	23	82.1	153	5	ABP26146	ABP26146 Streptoco	794	23	82.1	237	5	ABO01182	ABO01182 Human ORF
722	23	82.1	157	3	AB40581	Aab40581 Human ORF	795	23	82.1	239	5	ABG01182	Abg01182 Novel hum
723	23	82.1	157	5	ABP32349	ABP32349 Human his	796	23	82.1	240	4	ABG23621	Abg23621 Novel hum
724	23	82.1	161	5	AAE24343	AAE24343 Human lun	797	23	82.1	241	4	ABG07121	Abg07121 Novel hum
725	23	82.1	169	4	ABG07155	Abg07155 Novel hum	798	23	82.1	244	4	ABG05011	Abg05011 Novel hum
726	23	82.1	170	2	AAE21522	AAE21522 Protein e	799	23	82.1	244	4	ABG13384	Abg13384 Novel hum
727	23	82.1	170	2	AAV52782	AAV52782 Treponema	800	23	82.1	244	4	ABG13215	Abg13215 Novel hum
728	23	82.1	171	4	ABG07118	Abg07118 Novel hum	801	23	82.1	244	4	ABG07110	Abg07110 Novel hum
729	23	82.1	172	4	ABG07134	Abg07134 Novel hum	802	23	82.1	245	6	ABU23960	Abu23960 Protein e
730	23	82.1	174	5	ABP30514	ABP30514 Streptoco	803	23	82.1	246	4	ABG26936	Abg26936 Novel hum
731	23	82.1	174	7	ADG88448	ADG88448 Arabidops	804	23	82.1	246	4	ABG07129	Abg07129 Novel hum
732	23	82.1	175	4	ABG07124	Abg07124 Novel hum	805	23	82.1	247	4	ABG07114	Abg07114 Novel hum
733	23	82.1	178	3	AG45401	Aag45401 Arabidops	806	23	82.1	250	4	ABG07120	Abg07120 Novel hum
734	23	82.1	178	5	ABU51446	Abu51446 Helicobac	807	23	82.1	250	3	AG45400	Aag45400 Arabidops
735	23	82.1	179	6	ABU48693	ABU48693 Protein e	808	23	82.1	251	3	AG45400	Aag45400 Arabidops
736	23	82.1	180	6	ABU19238	ABU19238 Protein e	809	23	82.1	254	5	ABP73674	Abp73674 Candida a
737	23	82.1	181	7	ADF05683	ADF05683 Bacterial	810	23	82.1	254	6	ABU02851	Abu02851 S. pneumo
738	23	82.1	182	6	ABM69676	ABM69676 Photorhab	811	23	82.1	254	5	ABU02851	Abu02851 S. pneumo
739	23	82.1	183	2	AAV37028	AAV37028 Amino aci	812	23	82.1	256	6	ADK63524	ADK63524 Disease c
740	23	82.1	183	4	ABG04951	Abg04951 Novel hum	813	23	82.1	256	7	ADM25415	ADM25415 Hyperther
741	23	82.1	183	7	ABG070422	Abg070422 Pseudomon	814	23	82.1	257	4	ABM82206	ABM82206 Megalomic
742	23	82.1	184	4	ABG13438	Abg13438 Novel hum	815	23	82.1	257	4	ABM82206	ABM82206 Megalomic
743	23	82.1	184	6	ABM69672	ABM69672 Photorhab	816	23	82.1	258	6	ABM68629	ABM68629 Photorhab
744	23	82.1	186	2	AY0708067	AY0708067 Rat serin	817	23	82.1	258	6	ABU25033	Abu25033 Protein e
745	23	82.1	191	4	ABG07138	Abg07138 Novel hum	818	23	82.1	259	6	ABU25033	Abu25033 Protein e
746	23	82.1	192	5	ABG69479	ABG69479 Human bai	819	23	82.1	260	3	AAV75387	AAV75387 Neisseria
747	23	82.1	193	2	AAV08491	AAV08491 Rat serin	820	23	82.1	260	3	AAV75386	AAV75386 Neisseria
748	23	82.1	195	4	ABG07143	Abg07143 Novel hum	821	23	82.1	261	4	ABG07157	Abg07157 Murine ad
749	23	82.1	195	7	ADH88273	ADH88273 Enterococ	822	23	82.1	264	8	ADM67197	ADM67197 Murine ad
750	23	82.1	200	4	ABG01444	ABG01444 Novel hum	823	23	82.1	264	6	ABU70978	Abu70978 Human adi
751	23	82.1	201	4	ABG13598	Abg13598 Novel hum	824	23	82.1	268	6	ABM65529	ABM65529 Drosophil
752	23	82.1	203	5	ADK35061	ADK35061 Novel hum	825	23	82.1	269	4	ABU39270	Abu39270 Protein e
753	23	82.1	203	6	ADB11414	ADB11414 Alloiococ	826	23	82.1	270	6	ADC19708	Adc19708 H. influe
754	23	82.1	204	4	ABG07144	Abg07144 Novel hum	827	23	82.1	271	4	ABG07142	Abg07142 Novel hum
755	23	82.1	205	7	ADG88446	ADG88446 Arabidops	828	23	82.1	271	8	ADJ49765	ADJ49765 Oil-assoc



829	23	82.1	273	8	ADN46245	Adn46245	Thermococ	902	23	82.1	336	5	ABB53868	Abb53868	Lactococc
830	23	82.1	275	3	AAy49439	Aay49439	M. smegma	903	23	82.1	336	6	ADA34551	Ada34551	Acinetoba
831	23	82.1	275	7	ADG88445	Adg88445	Arabidops	904	23	82.1	337	4	ABG26934	Abg26934	Novel hum
832	23	82.1	275	7	ADG88454	Adg88454	Arabidops	905	23	82.1	338	4	ABG26948	Abg26948	Novel hum
833	23	82.1	277	2	AAW18875	Aaw18875	Babesia e	906	23	82.1	338	7	ABO63655	AbO63655	Klebsiell
834	23	82.1	277	7	ADG88440	Adg88440	Arabidops	907	23	82.1	339	6	ABU24630	Abu24630	Protein e
835	23	82.1	278	4	ABG07127	Abg07127	Novel hum	908	23	82.1	341	4	AAU33914	Aau33914	Staphyloc
836	23	82.1	280	4	ABG07148	Abg07148	Novel hum	909	23	82.1	341	5	ABU52197	Abu52197	Helicobac
837	23	82.1	282	4	AAU16221	Aau16221	Human nov	910	23	82.1	341	5	ABBA47518	Abba47518	Listeria
838	23	82.1	282	5	ABP29659	Abp29659	Streptoco	911	23	82.1	341	6	ABU32777	Abu32777	Protein e
839	23	82.1	282	5	ABP30157	Abp30157	Streptoco	912	23	82.1	341	7	ADH87639	Adh87639	Enterococ
840	23	82.1	282	6	ABU55290	Abu55290	Human nov	913	23	82.1	343	4	ABG26945	Abg26945	Novel hum
841	23	82.1	283	3	AGC47120	Aag47120	Arabidops	914	23	82.1	345	3	AAAG38499	Aag38499	Arabidops
842	23	82.1	283	3	AGC53788	Aag53788	Arabidops	915	23	82.1	345	6	ABU19761	Abu19761	Protein e
843	23	82.1	283	3	AAAG17825	Aag17825	Arabidops	916	23	82.1	345	7	ABO61789	AbO61789	Klebsiell
844	23	82.1	288	4	AAAM80259	Aam80259	Human pro	917	23	82.1	346	4	ABG99854	Abg99854	S. cinnam
845	23	82.1	289	4	ABG10671	Abg10671	Novel hum	918	23	82.1	346	4	ABG07125	Abg07125	Novel hum
846	23	82.1	289	4	ABG13675	Abg13675	Novel hum	919	23	82.1	346	4	ABG24571	Abg24571	Novel hum
847	23	82.1	290	5	ABP51267	Abp51267	Human MDD	920	23	82.1	346	4	ABG26933	Abg26933	Novel hum
848	23	82.1	292	6	ABU37915	Abu37915	Protein e	921	23	82.1	347	6	ABG244239	Abg244239	Protein e
849	23	82.1	293	2	AAW69419	Aaw69419	Protein e	922	23	82.1	347	6	ADB09524	Adb09524	Alloiooc
850	23	82.1	294	4	ABG07145	Abg07145	Novel hum	923	23	82.1	347	7	AAy97208	Aay97208	Campyloba
851	23	82.1	298	4	ABB66015	Abb66015	Drosophil	924	23	82.1	347	7	ABR56209	AbR56209	Mutant Mk
852	23	82.1	298	4	ABG23387	Abg23387	Novel hum	925	23	82.1	349	3	AAAG46268	Aag46268	Arabidops
853	23	82.1	298	6	ABU21348	Abu21348	Protein e	926	23	82.1	349	3	AAAG24691	Aag24691	Arabidops
854	23	82.1	298	8	ADL33453	Adl33453	Lolium pe	927	23	82.1	350	8	ADO61781	AdO61781	Transcrip
855	23	82.1	300	4	ABG20320	Abg20320	Novel hum	928	23	82.1	352	2	AAW21895	Aaw21895	Phenylala
856	23	82.1	300	4	ABG22452	Abg22452	Novel hum	929	23	82.1	352	3	AAy97208	Aay97208	Campyloba
857	23	82.1	300	4	ABG20883	Abg20883	Novel hum	930	23	82.1	352	4	AAU36772	Aau36772	Staphyloc
858	23	82.1	301	5	ABP25609	Abp25609	Streptoco	931	23	82.1	352	4	ABG26964	Abg26964	Novel hum
859	23	82.1	301	4	ABG22399	Abg22399	Novel hum	932	23	82.1	352	6	ABP77954	Abp77954	N. gonorr
860	23	82.1	302	2	AAy28803	Aay28803	Maize his	933	23	82.1	352	6	ABJ18498	Abj18498	Campyloba
861	23	82.1	302	5	ABE07518	AbE07518	Human dru	934	23	82.1	352	6	AAE35920	Aae35920	S. aureus
862	23	82.1	305	2	AAy28802	Aay28802	Maize his	935	23	82.1	352	6	ABU15946	Abu15946	Protein e
863	23	82.1	306	3	ADB80898	AdB80898	Bacillus	936	23	82.1	352	6	ABU43451	Abu43451	Protein e
864	23	82.1	308	7	ADG88439	Adg88439	Arabidops	937	23	82.1	352	6	ABM71336	Abm71336	Staphyloc
865	23	82.1	310	4	AAU34187	Aau34187	Staphyloc	938	23	82.1	352	7	ADC61377	Adc61377	Vanilla p
866	23	82.1	311	2	AAy28804	Aay28804	Maize his	939	23	82.1	354	3	AAAG07694	Aag07694	Arabidops
867	23	82.1	311	4	ABG07139	Abg07139	Novel hum	940	23	82.1	354	8	ADJ49761	Adj49761	Oil-assoc
868	23	82.1	311	6	ABM70166	Abm70166	Photorhab	941	23	82.1	355	2	AAW98505	Aaw98505	H. pylori
869	23	82.1	312	4	ABG26961	Abg26961	Novel hum	942	23	82.1	355	4	ABG07162	Abg07162	Novel hum
870	23	82.1	313	6	ABU28735	Abu28735	Protein e	943	23	82.1	356	6	ADA34971	Ada34971	Acinetoba
871	23	82.1	314	4	ABG04948	Abg04948	Novel hum	944	23	82.1	357	7	ADG88443	Adg88443	Arabidops
872	23	82.1	314	5	ABBA48063	Abba48063	Listeria	945	23	82.1	358	3	AAAG53786	Aag53786	Arabidops
873	23	82.1	314	7	ADG88452	Adg88452	Arabidops	946	23	82.1	358	3	AAAG17823	Aag17823	Arabidops
874	23	82.1	315	4	ABG26640	Abg26640	Novel hum	947	23	82.1	358	3	AAAG47118	Aag47118	Arabidops
875	23	82.1	315	6	ABR41516	AbR41516	Human DIT	948	23	82.1	358	4	AAU28020	Aau28020	Novel hum
876	23	82.1	315	7	ADC32741	Adc32741	Human nov	949	23	82.1	358	4	ABG28646	Abg28646	Novel hum
877	23	82.1	317	7	ADF05976	Adf05976	Bacterial	950	23	82.1	358	4	ABG26956	Abg26956	Novel hum
878	23	82.1	319	4	AAU17094	Aau17094	Novel sig	951	23	82.1	358	7	ABG07172	Abg07172	Novel hum
879	23	82.1	319	6	ABU19682	Abu19682	Protein e	952	23	82.1	358	8	ABO71808	AbO71808	Pseudomon
880	23	82.1	319	7	ADB93802	AdB93802	Human nov	953	23	82.1	358	8	ADN74409	Adn74409	Thale cre
881	23	82.1	320	4	ABG10658	Abg10658	Novel hum	954	23	82.1	359	3	AAAG21283	Aag21283	Arabidops
882	23	82.1	324	4	ABE96702	AbE96702	Putative	955	23	82.1	360	4	ABG07151	Abg07151	Novel hum
883	23	82.1	325	3	AAAG47119	Aag47119	Arabidops	956	23	82.1	360	7	ADC61397	Adc61397	Nicotiana
884	23	82.1	325	3	AAAG17824	Aag17824	Arabidops	957	23	82.1	360	7	ADG61395	Adg61395	Arabidops
885	23	82.1	325	3	AAAG53787	Aag53787	Arabidops	958	23	82.1	360	7	ADG61398	Adg61398	Arabidops
886	23	82.1	326	4	ABE61578	AbE61578	Drosophil	959	23	82.1	361	5	ABP39017	Abp39017	Staphyloc
887	23	82.1	327	6	ABU36370	Abu36370	Protein e	960	23	82.1	362	2	AAAR48694	Aar48694	G-protein
888	23	82.1	328	6	ABU35570	Abu35570	Protein e	961	23	82.1	362	2	AAW02666	Aaw02666	G-protein
889	23	82.1	329	7	ADH88353	Adh88353	Enterococ	962	23	82.1	362	2	AAEL1410	Aae1410	Beta-2 ad
890	23	82.1	331	6	ABU49759	Abu49759	Protein e	963	23	82.1	363	5	ABBA47702	Abb47702	Listeria
891	23	82.1	332	4	ABB57987	AbB57987	Drosophil	964	23	82.1	363	6	ABU32432	Abu32432	Protein e
892	23	82.1	333	3	AAAG21284	Aag21284	Arabidops	965	23	82.1	363	7	ADC61399	Adc61399	Zea mays
893	23	82.1	333	4	ABG07123	Abg07123	Novel hum	966	23	82.1	364	4	ABG26941	Abg26941	Novel hum
894	23	82.1	333	4	ABG26927	Abg26927	Novel hum	967	23	82.1	366	4	ABG207126	Abg207126	Novel hum
895	23	82.1	334	2	AAW82627	Aaw82627	Ehrlichia	968	23	82.1	368	4	ABG10027	Abg10027	Novel hum
896	23	82.1	334	3	AAW78541	Aay78541	Ehrlichia	969	23	82.1	368	6	ABP57456	Abp57456	Mycobacte
897	23	82.1	334	5	ABG93400	Abg93400	Ehrlichia	970	23	82.1	368	6	ABU35850	Abu35850	Protein e
898	23	82.1	334	5	AAU73197	Aau73197	Human gra	971	23	82.1	368	6	ABU36747	Abu36747	Protein e
899	23	82.1	335	3	AAy76008	Aay76008	Rat EGF e	972	23	82.1	368	6	ABU34797	Abu34797	Protein e
900	23	82.1	335	4	AAE55947	Aae55947	Skin cell	973	23	82.1	369	4	ABG07159	Abg07159	Novel hum
901	23	82.1	335	5	ABB72147	AbB72147	Rat prote	974	23	82.1	370	7	ADG88444	Adg88444	Arabidops

975 23 82.1 371 4 ABG26925  
976 23 82.1 371 4 ABG24561  
977 23 82.1 371 4 ABG07385  
978 23 82.1 371 6 ABU33790  
979 23 82.1 371 6 ABU25841  
980 23 82.1 372 7 ADA01478  
981 23 82.1 373 4 ABG26946  
982 23 82.1 373 5 ABH91084  
983 23 82.1 373 8 ADN74463  
984 23 82.1 373 8 ADP29741  
985 23 82.1 374 3 AAG24690  
986 23 82.1 374 3 AAG46267  
987 23 82.1 374 6 ABM67064  
988 23 82.1 377 4 ABG24555  
989 23 82.1 378 4 AAB65775  
990 23 82.1 378 4 ABG07158  
991 23 82.1 378 7 ADB94794  
992 23 82.1 379 4 ABG26952  
993 23 82.1 380 4 ABG07150  
994 23 82.1 384 4 ABG10303  
995 23 82.1 385 7 ADG98840  
996 23 82.1 385 8 ADO29655  
997 23 82.1 386 4 ABB61754  
998 23 82.1 386 4 ABG24562  
999 23 82.1 386 4 ABG07122  
1000 23 82.1 386 6 ABP57446

ALIGNMENTS

RESULT 1  
ID ADF70007 standard; protein; 45 AA.  
XX  
AC ADF70007;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE AcmA-type homologue amino acid sequence.  
XX  
KW delivery; targeting system; AcmA-type anchor protein; solid tumour;  
KW health; medical; agricultural; cosmetic; controlled release.  
XX  
OS Lactococcus lactis.  
XX  
PN WO2003084508-A1.  
XX  
PD 16-OCT-2003.  
XX  
PF 04-APR-2003; 2003WO-NL000256.  
XX  
PR 04-APR-2002; 2002EP-00076316.  
PR 04-APR-2002; 2002US-0369927P.  
PR 05-APR-2002; 2002US-0370485P.  
PR 20-DEC-2002; 2002EP-00080481.  
XX  
PA (NANO-) APPLIED NANOSYSTEMS BV.  
XX  
XX Friesen RHE, Leenhouts CJ, Hektor HJ, Van Esch JH, Heeres A;  
PI Robillard GT;  
XX  
DR WPI; 2003-877005/81.  
XX  
XX Vehicle for targeted delivery of therapeutic or diagnostic agents,  
PT includes protein anchor and system for inducing availability of the  
PT agent.  
XX  
XX Example 3; Page 191; 303pp; English.  
PS  
XX The present invention describes a vehicle (A) for delivering a substance  
CC (I) to a predetermined site, which comprises (I); a system for inducing  
CC availability of at least one compartment of (A) towards the exterior;  
CC

CC and, as targeting system for directing (A) to the site, an AcmA-type  
CC anchor protein (II). (A) are used for delivery of diagnostic and  
CC therapeutic agents to predetermined sites in the body, particularly  
CC joints or solid tumours but can be used more generally for health,  
CC medical, agricultural and cosmetic applications. (A) significantly  
CC increases the half-life of peptides in the circulation and, by providing  
CC controlled release, ensures relatively high bioavailability, allowing  
CC therapeutic use of agents that would otherwise be too toxic for systemic  
CC administration. The native AcmA peptide targets Gram-positive bacteria  
CC but its homologues can be engineered to have different selectivity. The  
CC present sequence is used in the exemplification of the present invention.  
SQ Sequence 45 AA;  
  
Query Match 96.4%; Score 27; DB 7; Length 45;  
Best Local Similarity 83.3%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 GYVVEE 6  
Db 18 GYSVEE 23  
  
RESULT 2  
ID AEM68208 standard; protein; 235 AA.  
XX  
AC AEM68208;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Photorhabdus luminescens protein sequence #1305.  
XX  
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
KW detection; food; gene expression; plant; animal; microorganism; toxin;  
KW antibiotic; biopesticide; virulence factor; disease model; plague;  
KW whooping cough.  
XX  
OS Photorhabdus luminescens.  
XX  
PN WO200294867-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 07-FEB-2002; 2002WO-IB003040.  
XX  
PR 07-FEB-2001; 2001FR-00001659.  
XX  
PA (INSP) INST PASTEUR.  
PA (CNRS) CNRS CENT NAT RECH SCI.  
XX  
XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
PI Buchrieser C;  
XX  
XX WPI; 2003-148459/14.  
XX  
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
XX  
XX Claim 2; SEQ ID NO 1305; 1205pp; French.  
XX  
XX The invention relates to the isolation of genes and their encoded  
CC proteins from Photorhabdus luminescens. The isolated sequences are  
CC sources of probes and primers for detecting the genome of P. luminescens  
CC and related species; to study polymorphisms; for gene analysis and for  
CC detection/amplification of the genes. Antibodies (Ab) raised against the  
CC polypeptides encoded by the genes are used for detection/identification  
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
CC carry a gene-containing vector are used to select compounds that  
CC modulate, regulate, induce or inhibit expression of the genes in plants,  
CC animals or microorganisms other than P. luminescens and are able to alter  
CC response or sensitivity to toxins and antibiotics produced by P.  
CC luminescens. Cells transformed to express the genes are useful for

CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically (to treat microbial infection by bacteria or fungi that  
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which P.  
 CC luminescens is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated P. luminescens proteins  
 XX  
 SQ Sequence 235 AA;

Query Match 96.4%; Score 27; DB 6; Length 235;  
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 |||||  
 DB 60 GYTVEE 65

## RESULT 3

AA13560  
 ID AAY13560 standard; protein; 282 AA.

XX  
 AC AAY13560;

XX  
 DT 30-JUL-1999 (first entry)

XX  
 DE S. haemolyticus D-amino acid aminotransferase.

XX  
 KW Listeria; vaccine; auxotrophic; attenuated; mutation; dat gene; dal gene;  
 KW T-cell response; HIV-1 antigen; HIV; aminotransferase.

XX  
 OS Staphylococcus haemolyticus.

XX  
 PN WO9925376-A1.

XX  
 PD 27-MAY-1999.

XX  
 PF 13-NOV-1998; 98WO-US024357.

XX  
 PR 18-NOV-1997; 97US-00972902.

XX  
 PA (UYPE-) UNIV PENNSYLVANIA.

XX  
 PI Frankel FR, Portnoy DA;

XX  
 DR WPI; 1999-337876/28.

XX  
 PT Attenuated Listeria strain which is auxotrophic for D-alanine synthesis.

XX  
 PS Disclosure; Fig 4A-B; 67pp; English.

XX The invention provides a vaccine comprising an auxotrophic attenuated  
 CC strain of Listeria which expresses an antigen, where the strain comprises  
 CC a mutation in at least one gene essential for growth of the Listeria  
 CC especially that the strain is auxotrophic for D-alanine. The gene is  
 CC selected from the Listeria dat and dal genes. The vaccine can be used in  
 CC mammals to stimulate T-cell responses to an antigen, especially an HIV-1  
 CC antigen. This is useful in creating effective vaccines especially against  
 CC HIV. The attenuated Listeria strain can be used as a bacterial vaccine  
 CC vector to raise an immunoreponse in e.g. immunocompromised patient or a  
 CC pregnant patient. The present sequence represents a deduced D-amino acid  
 CC aminotransferase of S. haemolyticus  
 XX

SQ Sequence 282 AA;

Query Match 96.4%; Score 27; DB 2; Length 282;

Best Local Similarity 83.3%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 |||||

DB 64 GYTVEE 69  
 |||||

## RESULT 4

AAU33563  
 ID AAU33563 standard; protein; 344 AA.

XX  
 AC AAU33563;

XX  
 DT 14-FEB-2002 (first entry)

XX  
 DE Pseudomonas aeruginosa cellular proliferation protein #7.

XX  
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 KW antibacterial; drug design.

XX  
 OS Pseudomonas aeruginosa.

XX  
 PN WO200170955-A2.

XX  
 PD 27-SEP-2001.

XX  
 PF 21-MAR-2001; 2001WO-US009180.

XX  
 PR 21-MAR-2000; 2000US-0191078P.

XX  
 PR 23-MAY-2000; 2000US-0206848P.

XX  
 PR 26-MAY-2000; 2000US-0207727P.

XX  
 PR 23-OCT-2000; 2000US-0242578P.

XX  
 PR 27-NOV-2000; 2000US-0253625P.

XX  
 PR 22-DEC-2000; 2000US-0257931P.

XX  
 PR 16-FEB-2001; 2001US-0259308P.

XX  
 PA (ELIT-) ELITRA PHARM INC.

XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;

XX  
 DR WPI; 2001-611495/70.

XX  
 DR N-PSDB; AAS51422.

XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.

XX  
 PS Example 3; SEQ ID NO 5059; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes,  
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence represents an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 344 AA;

Query Match 96.4%; Score 27; DB 4; Length 344;

Best Local Similarity 83.3%; Pred. No. 2.9e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 |||||

Db 322 GYAVEE 327

RESULT 5  
ABU15587  
ID ABU15587 standard; protein; 344 AA.  
XX  
XX  
AC ABU15587;  
XX  
XX 19-JUN-2003 (first entry)  
XX  
XX Protein encoded by Prokaryotic essential gene #1114.  
XX  
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
XX Pseudomonas aeruginosa.  
OS  
XX WO200277183-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
XX  
XX 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA19457.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 25; SEQ ID NO 43511; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC on a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at

ftp.wipo.int/pub/published\_pct\_sequences

Sequence 344 AA;  
Query Match 96.4%; Score 27; DB 6; Length 344;  
Best Local Similarity 83.3%; Pred. No. 2.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVEE 6  
Db 322 GYAVEE 327

RESULT 6  
ABO78798  
ID ABO78798 standard; protein; 349 AA.  
XX  
XX ABO78798;  
XX  
XX 29-JUL-2004 (first entry)  
XX  
XX Pseudomonas aeruginosa polypeptide #10973.  
XX  
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
OS  
XX Pseudomonas aeruginosa.  
XX US6551795-B1.  
XX  
XX 22-APR-2003.  
XX  
XX 18-FEB-1999; 99US-00252991.  
XX  
XX 18-FEB-1998; 98US-0074788P.  
PR 27-JUL-1998; 98US-0094190P.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
PI  
XX  
XX WPI; 2003-615309/58.  
DR N-PSDB; ABD12369.  
XX  
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
XX Disclosure; SEQ ID NO 27544; 455pp; English.  
XX  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html  
XX  
XX Sequence 349 AA;  
Query Match 96.4%; Score 27; DB 7; Length 349;  
Best Local Similarity 83.3%; Pred. No. 2.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVEE 6

```

Db      327 GYAVEE 332
      || |||
RESULT 7
ADC95552
ID  ADC95552 standard; protein; 354 AA.
XX
AC  ADC95552;
XX
DT  01-JAN-2004 (first entry)
XX
DE  E. faecium protein sequence SEQ ID 5179.
XX
KW  Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KW  abdominal-pelvic infection.
XX
OS  Enterococcus faecium.
XX
PN  US583275-B1.
XX
PD  24-JUN-2003.
XX
PF  30-JUN-1998; 98US-00107532.
XX
PR  02-JUL-1997; 97US-0051571P.
PR  14-MAY-1998; 98US-0085598P.
XX
PA  (GENO-) GENOME THERAPEUTICS CORP.
XX
PI  Doucette-Stamm LA, Bush D;
XX
WPI; 2003-799836/75.
DR  N-PSDB; ADC91898.
XX
PT  New isolated nucleic acid derived from Enterococcus faecium encoding an
PT  Enterococcus faecium polypeptide useful for detection, prevention and
PT  treatment of a pathological condition resulting from a bacterial
PT  infection.
XX
PS  Example 1; SEQ ID NO 5179; 243pp; English.
XX
CC  The invention relates to an isolated nucleic acid derived from
CC  Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC  one of 10 fully defined sequences given in the (or comprising 40
CC  sequential nucleotides chosen from any of the nucleic acids, its
CC  complement or sequences hybridising to it). Also included are a
CC  recombinant vector comprising the nucleic acid operably linked to
CC  transcription regulatory element, a cell comprising the vector and a
CC  single-stranded probe comprising the nucleic acid. The nucleic acids are
CC  chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC  The nucleic acids is useful for diagnosing pathological conditions
CC  resulting from E. faecium bacterial infection (e.g. urinary tract
CC  infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC  infection) and for screening drugs such as agonists and antagonists. The
CC  nucleic acid is useful for recombinant production of Candida albicans -
CC  derived peptides or antisense polypeptides. Pharmaceutical compositions
CC  and vaccines containing the nucleic acid are useful for preventing or
CC  treating Enterococcus faecium infections. The present sequence represents
CC  one if the disclosed E. faecium proteins.
XX
SQ  Sequence 354 AA;
      Query Match      96.4%; Score 27; DB 7; Length 354;
      Best Local Similarity 83.3%; Pred. No. 3e+02;
      Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
      || |||
Db      121 GYSVEE 126

RESULT 8
      Query Match      96.4%; Score 27; DB 7; Length 354;
      Best Local Similarity 83.3%; Pred. No. 3e+02;
      Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
      || |||
Db      121 GYSVEE 126

```

```

ADM25471
ID  ADM25471 standard; protein; 385 AA.
XX
AC  ADM25471;
XX
DT  20-MAY-2004 (first entry)
XX
DE  Hyperthermophile Methanopyrus kandleri protein #77.
XX
KW  hyperthermophile; protein stability enhancement;
KW  protein activity enhancement.
XX
OS  Methanopyrus kandleri.
XX
PN  WO2003076575-A2.
XX
PD  18-SEP-2003.
XX
PF  04-MAR-2003; 2003WO-US006664.
XX
PR  04-MAR-2002; 2002US-0361742P.
PR  14-MAY-2002; 2002US-0380423P.
PR  16-SEP-2002; 2002US-0410974P.
XX
PA  (FIDE-) FIDELITY SYSTEMS INC.
PA  (MALY/) MALYKH A.
XX
PI  Slesarev AI, Pavlov A, Pavlova N, Kozyavkin S;
XX
WPI; 2003-748383/70.
DR  N-PSDB; ADM27081.
XX
PT  New isolated nucleic acids encoding any of about 1700 Methanopyrus
PT  kandleri proteins, and the encoded proteins, useful as medicaments or
PT  as diagnostic agents.
XX
PS  Claim 31; SEQ ID NO 77; 1023pp; English.
XX
CC  The invention comprises the amino acid sequence of proteins from the
CC  hyperthermophile Methanopyrus kandleri, the invention also comprises the
CC  complete genome from Methanopyrus kandleri. The Methanopyrus kandleri
CC  proteins of the invention are useful for enhancing the stability and/or
CC  activity of other proteins. The Methanopyrus kandleri genome is useful in
CC  a variety of diagnostic and analytical methods. The present amino acid
CC  sequence represents a Methanopyrus kandleri protein of the invention.
XX
SQ  Sequence 385 AA;
      Query Match      96.4%; Score 27; DB 7; Length 385;
      Best Local Similarity 83.3%; Pred. No. 3.3e+02;
      Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
      || |||
Db      10 GYSVEE 15

RESULT 9
ABP26813
ID  ABP26813 standard; protein; 413 AA.
XX
AC  ABP26813;
XX
DT  02-JUL-2002 (first entry)
XX
DE  Streptococcus polypeptide SEQ ID NO 2802.
XX
KW  Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW  group A streptococcus; Streptococcus pyogenes; antibacterial;
KW  antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS  Streptococcus agalactiae.
XX

```

PN W0200234771-A2.  
 XX  
 PD 02-MAY-2002.  
 XX  
 PF 29-OCT-2001; 2001WO-GB004789.  
 XX  
 PR 27-OCT-2000; 2000GB-00026333.  
 PR 24-NOV-2000; 2000GB-00028727.  
 PR 07-MAR-2001; 2001GB-00005640.  
 XX  
 XX (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 XX Telford J, Massignani V, Margarit Y Rosl, Grandi G, Fraser C;  
 PI Tettelin H;  
 XX  
 XX WPI; 2002-352536/38.  
 DR N-PSDB; ABN67444.  
 XX  
 XX New Streptococcus protein for the treatment or prevention of infection or  
 PT disease caused by Streptococcus bacteria, such as meningitis, and for  
 PT detecting a compound that binds to the protein.  
 XX  
 XX Claim 1; Page 3431; 4525pp; English.  
 XX  
 XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 543 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins  
 XX  
 XX Sequence 413 AA;  
 SQ  
 Query Match 96.4%; Score 27; DB 5; Length 413;  
 Best Local Similarity 83.3%; Pred. No. 3.6e+02; Mismatches 1; Indels 0; Gaps 0;  
 Matches 5; Conservative 0;  
 QY 1 GYXVEE 6  
 Db 306 GYSVEE 311  
 XX  
 RESULT 10  
 ABB54231  
 ID ABB54231 standard; protein; 466 AA.  
 XX  
 AC ABB54231;  
 XX  
 XX 29-AUG-2003 (revised)  
 DT 16-MAY-2002 (first entry)  
 XX  
 XX Lactococcus lactis protein tagH.  
 DE  
 XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.  
 KW  
 XX Lactococcus lactis; IL1403.  
 OS  
 XX  
 XX FR2807446-A1.  
 PN  
 XX 12-OCT-2001.  
 PD  
 XX

PF 11-APR-2000; 2000FR-00004630.  
 XX  
 PR 11-APR-2000; 2000FR-00004630.  
 XX  
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.  
 XX  
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;  
 XX WPI; 2002-043418/06.  
 DR  
 XX New nucleotide sequence useful in the identification of Lactococcus  
 PT lactis and related species.  
 PT  
 XX Claim 6; SEQ ID NO 933; 2504pp; French.  
 XX  
 XX The present invention is related to a Lactococcus lactis nucleotide  
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic  
 CC acid sequence is useful in the detection and/or amplification of nucleic  
 CC acid sequence, particularly to identify Lactococcus lactis or related  
 CC species. The proteins of the invention are useful for the biosynthesis or  
 CC biodegradation of a composition of interest. The invention helps research  
 CC in lactic bacteria, particularly useful in the production of yogurt and  
 CC cheese. Note: The sequence data for this patent is based on equivalent  
 CC patent WO200177334 (published 19-OCT-2001) which is available in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 29-AUG-2003 to  
 CC standardise OS field)  
 CC  
 XX Sequence 466 AA;  
 SQ  
 Query Match 96.4%; Score 27; DB 5; Length 466;  
 Best Local Similarity 83.3%; Pred. No. 4.1e+02; Mismatches 1; Indels 0; Gaps 0;  
 Matches 5; Conservative 0;  
 QY 1 GYXVEE 6  
 Db 421 GYSVEE 426  
 XX  
 RESULT 11  
 AAU34095  
 ID AAU34095 standard; protein; 545 AA.  
 XX  
 AC AAU34095;  
 XX  
 XX 13-FEB-2002 (first entry)  
 DT  
 XX Staphylococcus aureus cellular proliferation protein #371.  
 DE  
 XX Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 KW antibacterial; drug design.  
 KW  
 XX Staphylococcus aureus.  
 OS  
 XX WO200170955-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX  
 XX 21-MAR-2001; 2001WO-US009180.  
 PF  
 XX 21-MAR-2000; 2000US-0191078P.  
 PR 23-MAY-2000; 2000US-0206848P.  
 PR 26-MAY-2000; 2000US-0207727P.  
 PR 23-OCT-2000; 2000US-0242578P.  
 PR 27-NOV-2000; 2000US-0253625P.  
 PR 22-DEC-2000; 2000US-0257931P.  
 PR 16-FEB-2001; 2001US-0269308P.  
 XX  
 XX (ELIT-) ELITRA PHARM INC.  
 PA  
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX

DR WPI; 2001-611495/70.  
 DR N-PSDB; AAS51954.  
 XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.  
 XX Example 3; SEQ ID NO 5591; 511pp; English.  
 XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes  
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence represents an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 545 AA;  
 SQ

Query Match 96.4%; Score 27; DB 4; Length 545;  
 Best Local Similarity 83.3%; Pred. No. 4.9e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVEE 6  
 |||||  
 Db 269 GYSVEE 274

RESULT 12  
 ID AU336605 standard; protein; 553 AA.  
 XX AU336605;  
 AC AU336605;  
 XX 14-FEB-2002 (first entry)  
 DT  
 XX Staphylococcus aureus cellular proliferation protein #775.  
 DE  
 XX Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 KW antibacterial; drug design.  
 KW  
 XX Staphylococcus aureus.  
 OS  
 XX WO200170955-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX 21-MAR-2001; 2001WO-US009180.  
 PF  
 XX 21-MAR-2000; 2000US-0191078P.  
 PR 23-MAY-2000; 2000US-0206848P.  
 PR 26-MAY-2000; 2000US-0207727P.  
 PR 23-OCT-2000; 2000US-0242578P.  
 PR 27-NOV-2000; 2000US-0253625P.  
 PR 22-DEC-2000; 2000US-0257931P.  
 PR 16-FEB-2001; 2001US-0269308P.  
 XX (ELIT-) ELITRA PHARM INC.  
 PA  
 XX Haselbeck R, Ohlseen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX WPI; 2001-611495/70.  
 DR

DR N-PSDB; AAS54464.  
 XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.  
 XX Example 3; SEQ ID NO 12198; 511pp; English.  
 XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes  
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence represents an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 553 AA;  
 SQ

Query Match 96.4%; Score 27; DB 4; Length 553;  
 Best Local Similarity 83.3%; Pred. No. 5e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVEE 6  
 |||||  
 Db 270 GYSVEE 275

RESULT 13  
 ID ABU16137 standard; protein; 553 AA.  
 XX ABU16137;  
 AC ABU16137;  
 XX 19-JUN-2003 (first entry)  
 DT  
 XX Protein encoded by prokaryotic essential gene #1664.  
 DE  
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 KW  
 XX Staphylococcus aureus.  
 OS  
 XX WO200277183-A2.  
 PN  
 XX 03-OCT-2002.  
 PD  
 XX 21-MAR-2002; 2002WO-US009107.  
 PF  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 PA  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR N-PSDB; ACA20007.  
 DR  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 DR

PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 44061; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 553 AA;

Query Match 96.4%; Score 27; DB 6; Length 553;  
 Best Local Similarity 83.3%; Pred. No. 5e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 270 GYSVEE 275

RESULT 14  
 ABM72931  
 ID ABM72931 standard; protein; 553 AA.

XX AC ABM72931;

XX DT 20-NOV-2003 (first entry)

XX DE Staphylococcus aureus protein #2171.

XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;  
 KW enzymatic assay; antibiotic target.

XX OS Staphylococcus aureus.

XX PN WO200294868-A2.

XX PD 28-NOV-2002.

XX PF 27-MAR-2002; 2002WO-IB002637.

XX PR 27-MAR-2001; 2001GB-00007661.

XX PA (CHIR-) CHIRON SPA.

XX PI Masignani V, Mora M, Scarselli M;

XX WPI; 2003-120786/11.  
 DR N-PSDB; ACF74491.

XX New Staphylococcus aureus protein, useful as a vaccine for treating or  
 PT preventing Staphylococcal infection, specifically an infection caused by  
 PT *S. aureus*, e.g. sepsis.

PS Claim 1; SEQ ID NO 4342; 49pp; English.

XX The invention relates to novel genes and encoded proteins from  
 CC Staphylococcus aureus. A composition comprising the *S. aureus* protein, a  
 CC nucleic acid encoding the protein, or an antibody to the protein, is  
 CC useful as a pharmaceutical, particularly as a vaccine for treating or  
 CC preventing infection due to Staphylococcus bacteria, specifically an  
 CC infection caused by *S. aureus*. The composition is particularly useful for  
 CC treating or preventing sepsis in a patient. The composition can also be  
 CC used for diagnostics. The protein is also used in an assay for enzymatic  
 CC studies and as a target for antibiotics. This sequence represents one of  
 CC the novel *S. aureus* proteins of the invention

XX Sequence 553 AA;

Query Match 96.4%; Score 27; DB 6; Length 553;  
 Best Local Similarity 83.3%; Pred. No. 5e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 270 GYSVEE 275

RESULT 15  
 ADJ49360  
 ID ADJ49360 standard; protein; 653 AA.

XX AC ADJ49360;

XX DT 06-MAY-2004 (first entry)

XX DE Oil-associated gene related protein #860.

XX oil-associated gene; transgenic; enhanced seed oil; vegetable oil.

XX OS Unidentified.

XX PN US2004025202-A1.

XX PD 05-FEB-2004.

XX PF 14-MAR-2003; 2003US-00389566.

XX PR 15-MAR-2002; 2002US-0365301P.

XX PR 26-JUN-2002; 2002US-0391786P.

XX PR 26-JUN-2002; 2002US-0392018P.

XX PA (LAUR/) LAURIE C C.

XX PA (RAVA/) RAVANELLO M.

XX PA (SAVA/) SAVAGE T.

XX PA (LEDE/) LEDEUX J R.

XX PA (ROGE/) ROGERS J A.

XX PI Laurie CC, Ravanello M, Savage T, Ledoux JR, Rogers JA;

XX WPI; 2004-142683/14.

XX Novel recombinant DNA construct comprising a promoter functional in  
 PT plants operably linked to an oil-associated gene for producing transgenic  
 PT plant seed.

XX Example 3; SEQ ID NO 1364; 22pp; English.

XX The invention relates to a recombinant DNA construct comprising a



CC promoter functional in plants operably linked to an oil-associated gene.  
CC The construct is useful for transgenic plant seed which has in its genome  
CC the construct, that is functional in the plant to transcribe the oil-  
CC associated gene. The transgenic plant seed grows into a plant having  
CC enhanced seed oil as compared to wild type. The construct is useful for  
CC producing hybrid maize seed. The transgenic plant seed is useful for  
CC producing vegetable oil. The present sequence represents the amino acid  
CC sequence of an oil-associated gene related protein.  
XX  
SQ Sequence 653 AA;

Query Match 96.4%; Score 27; DB 8; Length 653;  
Best Local Similarity 83.3%; Pred. No. 6e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 492 GYSVEE 497

RESULT 16  
AAY00939  
ID AAY00939 standard; protein; 762 AA.  
XX  
AC AAY00939;

XX 02-JUN-1999 (first entry)  
XX  
XX Desulfurococcus M11 TL DNA polymerase 29PY1 protein sequence.  
XX  
XX DNA polymerase; thermophilic bacteria; DNA synthesis.  
XX  
XX Desulfurococcus sp.

XX WO9907837-A1.  
XX 18-FEB-1999.  
XX  
XX 06-AUG-1998; 98WO-US017152.  
XX  
XX 06-AUG-1997; 97US-00907166.  
XX  
XX (DIVE-) DIVERSA INC.

XX Callen W, Mathur EJ;  
XX  
XX WPI; 1999-180490/15.  
XX N-PSDB; AAX27286.  
XX  
XX DNA polymerases from extremely thermophilic bacteria - useful for DNA  
XX synthesis.  
XX  
XX Claim 1; Fig 5; 72pp; English.

XX This sequence is a DNA polymerase of the invention, that was isolated  
XX from a thermophilic bacteria. The polymerases are used in DNA synthesis  
XX and as immunogens to raise antibodies (useful for affinity purification  
XX and to screen for related enzymes). Fragments of the DNA encoding the  
XX polymerases are used as probes to isolate related or full-length  
XX sequences and to produce the recombinant polymerases. The polymerases  
XX catalyze DNA synthesis by the addition of deoxynucleotides to the 3' end  
XX of a polynucleotide chain, using a complementary polynucleotide strand as  
XX a template. The polymerases have optimum activity at over 60 deg. C and  
XX can renature and regain activity after exposure to temperatures above 70  
XX deg. C  
XX  
SQ Sequence 762 AA;

Query Match 96.4%; Score 27; DB 2; Length 762;  
Best Local Similarity 83.3%; Pred. No. 7.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

Db 689 GYSVEE 694  
|||  
RESULT 17  
AAE22116  
ID AAE22116 standard; protein; 762 AA.  
XX  
AC AAE22116;  
XX  
XX 25-JUL-2002 (first entry)  
XX  
XX Desulfurococcus sp. DNA polymerase.  
XX  
XX DNA polymerase; thermostable; enzyme.  
XX  
XX Desulfurococcus sp.  
XX  
XX Key Location/Qualifiers  
XX Misc-difference 601  
XX /note= "Encoded by SCG"

XX WO200220735-A2.  
XX 14-MAR-2002.  
XX  
XX 06-SEP-2001; 2001WO-US028007.  
XX  
XX 06-SEP-2000; 2000US-00656309.  
XX (DIVE-) DIVERSA CORP.  
XX  
XX Callen W, Mathur EJ, Short JM;  
XX  
XX WPI; 2002-362247/39.  
XX N-PSDB; AAD35185.

XX New thermostable polymerase useful for sequencing DNA, amplifying double  
XX stranded DNA, or incorporating a non-natural nucleotide or a nucleotide  
XX analog into a DNA molecule.

XX Claim 42; Page 147-150; 161pp; English.

XX The invention relates to thermostable DNA polymerases having high  
XX temperature polymerase activity, such as those derived from Pyrobobus  
XX fumaria and nucleic acid molecules encoding such polymerases. Polymerases  
XX are useful for catalysing the formation or repair of a nucleic acid  
XX sequence and for modifying small molecules. They are also useful for  
XX sequencing DNA molecules, for preparing cDNA from mRNA, for amplifying  
XX double stranded DNA molecules and for incorporating non-natural  
XX nucleotides or nucleotide analogues into a DNA molecule. The present  
XX sequence is Desulfurococcus sp. DNA polymerase

XX Sequence 762 AA;

Query Match 96.4%; Score 27; DB 5; Length 762;  
Best Local Similarity 83.3%; Pred. No. 7.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 689 GYSVEE 694

RESULT 18  
ABR53197  
ID ABR53197 standard; protein; 858 AA.  
XX  
XX ABR53197;

XX 20-JUN-2003 (first entry)  
XX  
XX Protein sequence #SEQ ID 1259.

XX KW Multiprotein complex; eukaryote; drug target; diagnosis.  
 XX OS Saccharomyces cerevisiae.  
 XX PN EP1258494-A1.  
 XX PD 20-NOV-2002.  
 XX PF 20-DEC-2001; 2001EP-00130253.  
 XX PR 15-MAY-2001; 2001EP-00111774.  
 XX PA (CELL-) CELLZONE AG.  
 XX PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;  
 PI Marzioch M, Schultz JD, Superti-Furga GD;  
 DR WPI; 2003-250078/25.  
 DR N-PSDB; ACC61239.  
 XX New isolated protein complexes useful for diagnosing a disease or  
 PT disorder, or as a target for an active agent of a pharmaceutical,  
 PT preferably a drug target in the treatment or prevention of disease or  
 PT disorder.  
 XX Disclosure; SEQ ID NO 1259; 17pp + Sequence Listing; English.  
 XX The invention relates to multiprotein complexes from eukaryotes. Proteins  
 CC of the invention and DNA sequences encoding them are given in records  
 CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are  
 CC obtainable by using a protein as a bait and isolating the set of proteins  
 CC which is attached thereto from cells. Such protein complexes may comprise  
 CC up to 30 distinct proteins. Protein complexes of the invention are useful  
 CC for diagnosing a disease or disorder, or as a target for an active agent  
 CC of a pharmaceutical, preferably a drug target in the treatment or  
 CC prevention of a disease or disorder. Note: The sequence data for this  
 CC patent is not represented in the printed specification, but is based on  
 CC sequence information supplied by the European Patent Office. The complete  
 CC document is available on CD-ROM  
 XX Sequence 858 AA;  
 Query Match 96.4%; Score 27; DB 6; Length 858;  
 Best Local Similarity 83.3%; Pred. No. 8.2e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVEE 6  
 DB 565 GYTVVE 570  
 RESULT 19  
 ADK63242  
 ID ADK63242 standard; protein; 858 AA.  
 AC ADK63242;  
 DT 06-MAY-2004 (first entry)  
 DE Disease treating protein complex-derived protein #747.  
 XX protein complex; drug target; diagnosis.  
 XX Unidentified.  
 XX EF1338608-A2.  
 XX 27-AUG-2003.  
 XX 20-DEC-2002; 2002EP-00102902.  
 XX 20-DEC-2001; 2001EP-00130253.  
 PR

XX (CELL-) CELLZONE AG.  
 XX Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;  
 PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;  
 PI Michon A, Leutwein C, Rick J;  
 XX WPI; 2003-638460/61.  
 DR N-PSDB; ADK63243.  
 XX New proteins and protein complexes from eukaryotes, useful as targets in  
 PT drug screening, or in diagnosing or screening for the presence of a  
 PT disease or disorder, or a predisposition for developing a disease or  
 PT disorder in a subject.  
 XX Disclosure; SEQ ID NO 1493; 13pp; English.  
 XX The invention relates to novel protein complexes comprising a first and a  
 CC second protein, or its derivative, fragment, homologue or variant. The  
 CC proteins are selected from given protein complexes, which are not defined  
 CC in the specification. The variants are encoded by nucleic acids that  
 CC hybridize to the nucleic acids encoding the proteins under low stringency  
 CC conditions. The protein complexes are useful as targets for an active  
 CC agent of a pharmaceutical. These protein complexes are particularly  
 CC useful as drugs targets for the treatment or preventing of a disease or  
 CC disorder. The complexes and methods above are useful in diagnosing or  
 CC screening for the presence of a disease or disorder or a predisposition  
 CC for developing a disease or disorder in a subject. These are also useful  
 CC in screening for a drug for treatment or prevention of a disease or  
 CC disorder. The molecule that modulates the amount, activity or protein  
 CC components of the complex is useful for the manufacture of a medicament  
 CC for the treatment or prevention of a disease or disorder. This sequence  
 CC corresponds to a protein of the invention. (Note: the sequence data for  
 CC this patent did not form part of the printed specification but was  
 CC obtained from the EPO in electronic format).  
 XX Sequence 858 AA;  
 Query Match 96.4%; Score 27; DB 7; Length 858;  
 Best Local Similarity 83.3%; Pred. No. 8.2e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVEE 6  
 DB 565 GYTVVE 570  
 RESULT 20  
 ABB77436  
 ID ABB77436 standard; protein; 1075 AA.  
 AC ABB77436;  
 DT 19-JUL-2002 (first entry)  
 DE Human tumour marker protein se33-1.  
 XX Human; tumour; cytostatic; cutaneous T cell lymphoma; CTL; vaccine;  
 KW antigen-presenting cell; tumour-specific T cell.  
 XX Homo sapiens.  
 OS WO200238803-A2.  
 PN 16-MAY-2002.  
 PD 08-NOV-2001; 2001WO-DE004229.  
 PF 08-NOV-2000; 2000DE-01055285.  
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 PA Eichmueller S, Schadendorf D, Usener D;  
 XX





XX Claim 1; Page 8; 28pp; French.

XX The present sequence represents a fragment of a mutant plant protein of

CC the GRAS family. The GRAS protein is encoded by a BZH gene. The

CC specification describes a mutant allele of the BZH gene. GRAS proteins

CC are transcription factors implicated in regulation of the response to

CC gibberellins and thus in control of morphogenesis and plant development.

CC The mutant GRAS protein is used to produce dwarf plants, specifically

CC crucifers. Dwarf plants may be sown earlier (increasing nitrate

CC accumulation without risking excessive stem growth during winter), and

CC have better resistance to cold and lodging. They are also easier to

XX harvest and allow for better monitoring of the crop

XX Sequence 6 AA;

Query Match 92.9%; Score 26; DB 4; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.7e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYXVEE 6

Db |||||

1 GYXVEE 6

RESULT 25

AAAR43265

ID AAR43265 standard; protein; 53 AA.

XX

AC AAR43265;

XX

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 04-MAY-1994 (first entry)

XX

DE RGA1 C-terminal.

XX

KW Rhamnogalacturonan acetyl esterase; RGAE; enzyme; Aspergillus aculeatus;

KW modified hairy region; MHR; probe; pYES 2.0; amplification; primer;

KW polymerase chain reaction; PCR.

XX

OS Aspergillus aculeatus; RE4.

XX

FN WO9320190-A1.

XX

PD 14-OCT-1993.

XX

PF 29-MAR-1993; 93WO-DK000109.

XX

PR 27-MAR-1992; 92DK-00000420.

XX

PA (NOVO ) NOVO-NORDISK AS.

XX

PI Doerreich K, Christensen FM, Schnell Y, Mischler M, Dalboge H;

PI Heldt-Hansen HP;

XX

DR WPI; 1993-336898/42.

DR N-PSDB; AAQ50196.

XX

XX Rhamno-galacturonan acetyl esterase enzyme obtd. from Aspergillus

PT aculeatus - used to degrade acetylated modified hairy region of vegetable

PT material.

XX

PS Disclosure; Fig 9; 58pp; English.

XX

CC The probe and primers (AAQ50192-94) were used in the amplification of

CC RGAE DNA from Aspergillus aculeatus RE4. The 5' and 3'-end sequences are

CC given in AAQ50195-96. (Updated on 25-MAR-2003 to correct PN field.)

CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 24-OCT-2003 to

CC standardise OS field)

XX

XX Sequence 53 AA;

Query Match 92.9%; Score 26; DB 4; Length 53;

Best Local Similarity 83.3%; Pred. No. 62;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6

Db |||||

41 GYXVEE 46

RESULT 26

AAAG99655

ID AAG99655 standard; peptide; 54 AA.

XX

AC AAG99655;

XX

DT 27-SEP-2001 (first entry)

XX

DE ERA binding domain polypeptide SEQ ID NO 97.

XX

KW ERA binding domain; Escherichia coli; GTPase; antimicrobial;

KW antibacterial; antibiotic; pathogenesis; infection; vaccine;

KW peptide therapy.

XX

OS Caenorhabditis elegans.

XX

PN WO200153458-A2.

XX

PD 26-JUL-2001.

XX

PF 17-JAN-2001; 2001WO-US001786.

XX

PR 18-JAN-2000; 2000US-0176870P.

XX

PA (SMIK ) SMITHKLINE BEECHAM CORP.

PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX

PI Lupas AN, Pearce KH;

XX

DR WPI; 2001-476108/51.

XX

XX New ERA binding domain polypeptides and polynucleotides encoding them,

PT useful as research reagents and materials for discovery of treatments and

PT diagnostics for diseases, or for genetic immunization.

XX

PS Claim 1; Page 12; 279pp; English.

XX

CC The present invention relates to ERA binding domain polypeptides

CC (AAG99559-AAG99989 and AAQ00010-AAQ00021). The era gene in Escherichia

CC coli codes for an essential GTPase protein able to autophosphorylate at

CC serine and/or threonine residues. The protein has potential antimicrobial

CC and antibacterial activity and is useful in screening for antagonists,

CC agonists and for compounds with antibiotic activity. The proteins are

CC also useful in determining their role in pathogenesis of infection,

CC dysfunction and disease and could be used as part of a vaccine and/or

CC peptide therapy

XX

XX Sequence 54 AA;

Query Match 92.9%; Score 26; DB 4; Length 54;

Best Local Similarity 83.3%; Pred. No. 64;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6

Db |||||

49 GYXVEE 54

RESULT 27

ADM26846

ID ADM26846 standard; protein; 64 AA.

XX

XX ADM26846;

XX



DR WPI; 2001-483447/52.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human fetal liver.  
XX Claim 27; SEQ ID NO 36323; 639pp + Sequence Listing; English.  
PS  
XX The invention relates to a single exon nucleic acid probe for measuring  
CC human gene expression in a sample derived from human foetal liver. The  
CC single exon nucleic acid probes may be used for predicting, measuring and  
CC displaying gene expression in samples derived from human fetal liver. The  
CC present sequence is a peptide encoded by a single exon nucleic acid probe  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 84 AA;  
Query Match 92.9%; Score 26; DB 4; Length 84;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GYXVEE 6  
Db 13 GYLVEE 18  
RESULT 30  
ID AAM37583 standard; protein; 84 AA.  
XX  
AC AAM37583;  
XX  
XX 17-OCT-2001 (first entry)  
DT  
XX Peptide #11620 encoded by probe for measuring placental gene expression.  
DE  
XX Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.  
KW  
OS Homo sapiens.  
XX  
XX WO200157272-A2.  
PN  
XX 09-AUG-2001.  
PD  
XX 30-JAN-2001; 2001WO-US000663.  
PF  
XX 04-FEB-2000; 2000US-0180312P.  
PR  
XX 26-MAY-2000; 2000US-0207456P.  
PR  
XX 30-JUN-2000; 2000US-00608408.  
PR  
XX 03-AUG-2000; 2000US-00632366.  
PR  
XX 21-SEP-2000; 2000US-0234687P.  
PR  
XX 27-SEP-2000; 2000US-0236359P.  
PR  
XX 04-OCT-2000; 2000GB-00024253.  
PR  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2001-488897/53.  
DR  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human placenta.  
PT  
XX Claim 27; SEQ ID NO 37852; 654pp; English.  
PS  
XX The present invention relates to single exon nucleic acid probes (SENPs;  
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders

XX  
SQ Sequence 84 AA;  
Query Match 92.9%; Score 26; DB 4; Length 84;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GYXVEE 6  
Db 13 GYLVEE 18  
RESULT 31  
ID ABB26636 standard; protein; 84 AA.  
XX  
AC ABB26636;  
XX  
XX 23-JAN-2002 (first entry)  
DT  
XX Protein #8635 encoded by probe for measuring heart cell gene expression.  
DE  
XX Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX  
OS Homo sapiens.  
XX  
XX WO200157274-A2.  
PN  
XX 09-AUG-2001.  
PD  
XX 30-JAN-2001; 2001WO-US000666.  
PF  
XX 04-FEB-2000; 2000US-0180312P.  
PR  
XX 26-MAY-2000; 2000US-0207456P.  
PR  
XX 30-JUN-2000; 2000US-00608408.  
PR  
XX 03-AUG-2000; 2000US-00632366.  
PR  
XX 21-SEP-2000; 2000US-0234687P.  
PR  
XX 27-SEP-2000; 2000US-0236359P.  
PR  
XX 04-OCT-2000; 2000GB-00024253.  
PR  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2001-488899/53.  
DR  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts.  
PT  
XX Claim 15; SEQ ID NO 28406; 530pp; English.  
PS  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 84 AA;  
Query Match 92.9%; Score 26; DB 4; Length 84;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GYXVEE 6

```
Db      || |||
      13 GYLVEE 18

RESULT 32
AAM64647
ID AAM64647 standard; protein; 84 AA.
XX
AC
XX
XX
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 36752.
XX
XX Human brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 36752; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention
XX
XX Sequence 84 AA;
Query Match 92.9%; Score 26; DB 4; Length 84;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 13 GYLVEE 18

RESULT 33
ABG59061
ID ABG59061 standard; peptide; 84 AA.
XX
XX AC
XX ABG59061;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver peptide, SEQ ID No 37709.
XX
XX
```

```
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 27; SEQ ID NO 37709; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABG47348-ABG59930 represent human
XX liver single exon encoded peptides of the invention. Note: The sequence
XX information for this patent does not appear in the printed specification
XX but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 84 AA;
Query Match 92.9%; Score 26; DB 4; Length 84;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 13 GYLVEE 18

RESULT 34
AAM21427
ID AAM21427 standard; protein; 90 AA.
XX
XX AC
XX AAM21427;
XX
XX 12-OCT-2001 (first entry)
XX
XX Peptide #7861 encoded by probe for measuring cervical gene expression.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
XX Homo sapiens.
XX
XX WO200157278-A2.
XX
```



PD	09-AUG-2001.	
XX		
XX	30-JAN-2001; 2001WO-US000670.	
XX		
PR	04-FEB-2000; 2000US-0180312P.	
PR	26-MAY-2000; 2000US-0207456P.	
PR	30-JUN-2000; 2000US-00608408.	
PR	03-AUG-2000; 2000US-00632366.	
PR	21-SEP-2000; 2000US-0234687P.	
PR	27-SEP-2000; 2000US-0236359P.	
PR	04-OCT-2000; 2000GB-00024263.	
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI; 2001-488901/53.	
XX		
PT	Human genome-derived single exon nucleic acid probes useful for analyzing	
PT	gene expression in human cervical epithelial cells.	
XX		
PS	Claim 27; SEQ ID NO 26253; 487pp; English.	
XX		
CC	The present invention relates to human single exon nucleic acid probes	
CC	(SENP: see Aa10068-AA128459). The present sequence is a peptide encoded	
CC	by one such probe. The SENPs are derived from human HeLa cells. The SENPs	
CC	can be used to produce a single exon microarray, which can be used for	
CC	measuring human gene expression in a sample derived from human cervical	
CC	epithelial cells. By measuring gene expression, the probes are therefore	
CC	useful in grading and/or staging of diseases of the cervix, notably	
CC	cervical cancer. Note: The sequence data for this patent did not form	
CC	part of the printed specification, but was obtained in electronic format	
CC	directly from WIPO at ftp.wipo.int/pub/published_pt_sequences	
XX		
SQ	Sequence 90 AA;	
Query Match	92.9%; Score 26; DB 4; Length 90;	
Best Local Similarity	83.3%; Pred. No. 1.1e+02;	
Matches	5; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
Qy	1 GYXVEE 6	
Db	49 GYLVEE 54	
RESULT 35		
ABB43765		
ID	ABB43765 standard; peptide; 90 AA.	
XX		
AC	ABB43765;	
XX		
DT	04-FEB-2002 (first entry)	
XX		
DE	Peptide #11271 encoded by human foetal liver single exon probe.	
XX		
KW	Human; foetal liver; gene expression; single exon nucleic acid probe.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157277-A2.	
XX		
PD	09-AUG-2001.	
XX		
PP	30-JAN-2001; 2001WO-US000669.	
XX		
PR	04-FEB-2000; 2000US-0180312P.	
PR	26-MAY-2000; 2000US-0207456P.	
PR	30-JUN-2000; 2000US-00608408.	
PR	03-AUG-2000; 2000US-00632366.	
PR	21-SEP-2000; 2000US-0234687P.	
PR	27-SEP-2000; 2000US-0236359P.	
PR	04-OCT-2000; 2000GB-00024263.	
XX		

CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders  
 XX  
 XX  
 SQ Sequence 90 AA;

Query Match 92.9%; Score 26; DB 4; Length 90;  
 Best Local Similarity 83.3%; Pred. No. 1.1e+02; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 1;

QY 1 GYXVEE 6  
 |||||  
 Db 49 GYLVEE 54

RESULT 37  
 ABB26708  
 ID ABB26708 standard; protein; 90 AA.

XX ABB26708;

AC ABB26708;

DT 23-JAN-2002 (first entry)

DE Protein #8707 encoded by probe for measuring heart cell gene expression.  
 XX Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000666.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human  
 XX hearts.  
 XX  
 XX Claim 15; SEQ ID NO 28478; 530pp; English.  
 XX  
 XX The present invention relates to single exon nucleic acid probes for  
 XX measuring human gene expression in a sample derived from human heart (see  
 XX ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 XX probe. The probes may be used for predicting, measuring and displaying  
 XX gene expression in samples derived from the human heart via microarrays.  
 XX By measuring gene expression, the probes are useful for predicting,  
 XX diagnosing, grading, staging, monitoring and prognosing diseases of the  
 XX human heart and vascular system e.g. cardiovascular disease, Note: The  
 XX hypertension, cardiac arrhythmias and congenital heart disease. sequence  
 XX data for this patent did not form part of the printed  
 XX specification, but was obtained in electronic format directly from WIPO  
 XX at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 90 AA;

XX Query Match 92.9%; Score 26; DB 4; Length 90;

Best Local Similarity 83.3%; Pred. No. 1.1e+02; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 1;

QY 1 GYXVEE 6  
 |||||  
 Db 49 GYLVEE 54

RESULT 38  
 AAM77498  
 ID AAM77498 standard; protein; 90 AA.

XX AAM77498;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 37804.

XX Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000668.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 XX gene expression in human bone marrow.  
 XX  
 XX Example 4; SEQ ID NO 37804; 658pp + Sequence Listing; English.  
 XX  
 XX The present invention provides a number of single exon nucleic acid  
 XX probes which are derived from genomic sequences expressed in the human  
 XX bone marrow. They can be used to measure gene expression in bone marrow  
 XX samples, which may enable the improved diagnosis and treatment of cancers  
 XX such as lymphoma, leukaemia and myeloma. The present sequence is a  
 XX protein encoded by one of the probes of the invention

XX Sequence 90 AA;

Query Match 92.9%; Score 26; DB 4; Length 90;  
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 |||||  
 Db 49 GYLVEE 54

RESULT 39  
 AAM64727  
 ID AAM64727 standard; protein; 90 AA.

XX AAM64727;

XX 05-NOV-2001 (first entry)

```
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 36832.
DE
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
KW
OS Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 36832; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention
XX
XX Sequence 90 AA;
XX
XX Query Match 92.9%; Score 26; DB 4; Length 90;
XX Best Local Similarity 83.3%; Pred. No. 1.1e+02;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GYXVEE 6
XX DB 49 GYLVEE 54
XX
XX RESULT 40
XX ABG59136
XX ID ABG59136 standard; peptide; 90 AA.
XX
XX AC ABG59136;
XX
XX DT 25-FEB-2003 (first entry)
XX
XX DE Human liver peptide, SEQ ID No 37784.
XX
XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX hypercholesterolaemia; coronary heart disease.
XX
XX OS Homo sapiens.
XX
XX WO200157273-A2.
XX
XX PN 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 36832.
DE
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
KW
OS Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 27; SEQ ID NO 37784; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABG47348-ABG59930 represent human
XX liver single exon encoded peptides of the invention. Note: The sequence
XX information for this patent does not appear in the printed specification
XX but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 90 AA;
XX
XX Query Match 92.9%; Score 26; DB 4; Length 90;
XX Best Local Similarity 83.3%; Pred. No. 1.1e+02;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GYXVEE 6
XX DB 49 GYLVEE 54
XX
XX RESULT 41
XX AAW78901
XX ID AAW78901 standard; protein; 102 AA.
XX
XX AC AAW78901;
XX
XX DT 25-MAR-2003 (revised)
XX 21-DEC-1998 (first entry)
XX
XX DE Human UNC-5 homologue UNC5H-2.
XX
XX KW UNC-5; UNC5H-1; human; netrin receptor; cell migration; axon guidance;
XX diagnosis; therapy.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX Misc-difference 11
XX /note= "encoded by TG"
XX
XX FT Misc-difference 19
XX /note= "encoded by AC"
XX
XX FT Misc-difference 48
XX /note= "encoded by TC"
XX
XX WO9837085-A1.
XX
XX 27-AUG-1998.
XX
```

PF 19-FEB-1998; 98WO-US003143.  
 XX  
 PR 19-FEB-1997; 97US-00808982.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;  
 XX  
 DR WPI; 1998-495364/42.  
 DR N-PSDB; AAV52943.  
 XX  
 XX Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and  
 PT the biopharmaceutical industry.  
 XX  
 XX Claim 1; Page 27; 32pp; English.  
 PS  
 XX UNC5H-2 and UNC5H-1 (see AAW78899) are human homologues of Caenorhabditis  
 CC elegans UNC-5 protein. Their amino acid sequences were deduced from  
 CC isolated unc5h cDNA clones (see AAV52943 and AAV52941) isolated from an  
 CC embryonic brain cDNA library. The predicted proteins show similarity with  
 CC UNC-5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin  
 CC type-1 repeats, a predicted membrane spanning region, and a large  
 CC intracellular domain. They are predicted to be involved in cell migration  
 CC and axon guidance, and are characterised as receptor proteins for  
 CC netrins. Rat UNC5H-1 (see AAW78898) and UNC5H-2 (see AAW78900) proteins  
 CC are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly  
 CC from transfected host cells. The invention also provides unc-5  
 CC hybridisation probes and primers, vertebrate UNC-5-specific binding  
 CC agents such as specific antibodies, and methods of making and using the  
 CC subject compositions in diagnosis (e.g. genetic hybridisation screens for  
 CC vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate  
 CC vertebrate unc-5 gene expression) and in the biopharmaceutical industry  
 CC (e.g. as immunogens, reagents for modulating cell guidance, reagents for  
 CC screening chemical libraries for lead pharmacological agents, etc.).  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 XX Sequence 102 AA;  
 Query Match 92.9%; Score 26; DB 2; Length 102;  
 Best Local Similarity 83.3%; Pred. NO. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVEE 6  
 DB 84 GYLVEE 89  
 |||||  
 |||||  
 RESULT 42  
 ADK16101  
 ID ADK16101 standard; protein; 102 AA.  
 XX  
 AC ADK16101;  
 XX  
 XX 06-MAY-2004 (first entry)  
 DT  
 XX  
 DE Nanoarchaeum equitans cancer-associated (CA) protein #26.  
 XX  
 XX cancer-associated gene; CA gene; cancer; carcinoma; lymphoma; leukaemia.  
 KW  
 XX Nanoarchaeum equitans.  
 OS  
 XX WO2003093434-A2.  
 XX  
 XX 13-NOV-2003.  
 PD  
 XX  
 XX 01-MAY-2003; 2003WO-US013699.  
 PF  
 XX  
 XX 01-MAY-2002; 2002US-0377447P.  
 PR  
 XX  
 XX (DIVE-) DIVERSA CORP.  
 PA  
 XX Stetter KO, Waters E, Kretz K, Podar M, Richardson T;  
 PI Noordewier M;  
 PI

XX WPI; 2004-053041/05.  
 DR N-PSDB; ADK16100.  
 XX  
 XX New recombinant cancer-associated genes, such as KCN39, useful for  
 PT diagnosing or treating carcinoma, e.g. breast, colon, rectal, pancreatic,  
 PT cervical, or skin cancers, lymphomas, or leukemia.  
 XX  
 PS Claim 64; SEQ ID NO 53; 251pp; English.  
 XX  
 CC The invention comprises then amino acid and coding sequences of cancer-  
 CC associated (CA) genes isolated from Nanoarchaeum equitans. The invention  
 CC also comprises the Nanoarchaeum equitans genome. The DNA and protein  
 CC sequences of the invention are useful for diagnosing and treating cancer  
 CC (e.g. carcinoma, lymphoma, or leukaemia). The present amino acid sequence  
 CC represents a Nanoarchaeum equitans CA protein of the invention.  
 XX  
 XX Sequence 102 AA;  
 Query Match 92.9%; Score 26; DB 8; Length 102;  
 Best Local Similarity 66.7%; Pred. NO. 1.3e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVEE 6  
 DB 50 GYTIEE 55  
 |||||  
 |||||  
 RESULT 43  
 ADC96202  
 ID ADC96202 standard; protein; 107 AA.  
 XX  
 AC ADC96202;  
 XX  
 XX 01-JAN-2004 (first entry)  
 DT  
 DE E. faecium protein sequence SEQ ID 5829.  
 XX  
 XX Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;  
 KW abdominal-pelvic infection.  
 XX  
 XX Enterococcus faecium.  
 OS  
 XX US6583275-B1.  
 PN  
 XX 24-JUN-2003.  
 PD  
 XX 30-JUN-1998; 98US-00107532.  
 PF  
 XX 02-JUL-1997; 97US-0051571P.  
 PR  
 XX 14-MAY-1998; 98US-0085598P.  
 PR  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 PA  
 XX Doucette-Stamm LA, Bush D;  
 PI  
 XX WPI; 2003-799836/75.  
 DR N-PSDB; ADC92548.  
 DR  
 XX New isolated nucleic acid derived from Enterococcus faecium encoding an  
 PT Enterococcus faecium polypeptide useful for detection, prevention and  
 PT treatment of a pathological condition resulting from a bacterial  
 PT infection.  
 XX  
 XX Example 1; SEQ ID NO 5829; 243pp; English.  
 PS  
 XX The invention relates to an isolated nucleic acid derived from  
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having  
 CC one of 10 fully defined sequences given in the (or comprising 40  
 CC sequential nucleotides chosen from any of the nucleic acids, its  
 CC complement or sequences hybridising to it). Also included are a  
 CC recombinant vector comprising the nucleic acid operably linked to a  
 CC transcription regulatory element, a cell comprising the vector and a

CC single-stranded probe comprising the nucleic acid. The nucleic acids are  
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.  
 CC The nucleic acids is useful for diagnosing pathological conditions  
 CC resulting from *E. faecium* bacterial infection (e.g. urinary tract  
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic  
 CC infection) and for screening drugs such as agonists and antagonists. The  
 CC nucleic acid is useful for recombinant production of *Candida albicans* -  
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions  
 CC and vaccines containing the nucleic acid are useful for preventing or  
 CC treating *Enterococcus faecium* infections. The present sequence represents  
 CC one if the disclosed *E. faecium* proteins.  
 XX

SQ Sequence 107 AA;

Query Match 92.9%; Score 26; DB 7; Length 107;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 || : ||  
 DB 86 GYSIEE 91

RESULT 44

AAB41337  
 ID AAB41337 standard; protein; 112 AA.

AC AAB41337;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF1101 polypeptide sequence SEQ ID NO:2202.

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnary; antiposoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.

XX Homo sapiens.

OS WO200058473-A2.

PN 05-OCT-2000.

PD 31-MAR-2000; 2000WO-US008621.

PR 31-MAR-1999; 99US-0127607P.

PR 02-APR-1999; 99US-0127636P.

PR 05-APR-1999; 99US-0127728P.

PR 30-MAR-2000; 2000US-00540763.

PA (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI; 2000-602362/57.

DR N-PSDB; AAC75546.

XX Novel nucleic acids and peptides derived from open reading frame X,

PT useful for treating e.g. cancers, proliferative disorders,

PT neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 1627; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 CC antiposoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX

SQ Sequence 112 AA;

Query Match 92.9%; Score 26; DB 3; Length 112;  
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 || : ||  
 DB 26 GYVEE 31

RESULT 45

ABP32756  
 ID ABP32756 standard; protein; 112 AA.

XX ABP32756;

AC ABP32756;

DT 08-JUL-2002 (first entry)

DE Human isomerase-like ORF1729 protein, SEQ ID NO:3458.

KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;  
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;  
 KW immune modulation; haematopoiesis regulation; tissue growth;  
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;  
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;  
 KW behaviour; cancer; proliferative disorder; neurological disorder;  
 KW cardiovascular disease; immune system disorder; organ transplantation;  
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnary;  
 KW vasotropic; antiposoriatic; antidiabetic; cytostatic; nootropic;  
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;  
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

OS Homo sapiens.

XX WO200190366-A2.

PN 29-NOV-2001.

PD 24-MAY-2001; 2001WO-US017076.

PR 24-MAY-2000; 2000US-0206690P.

PA (CURA-) CURAGEN CORP.

XX Leach MD, Shinkets RA;

XX WPI; 2002-106200/14.

DR N-PSDB; ABN76782.

XX Novel human polypeptides and polynucleotides useful for diagnosing,  
PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and disorders related to organ  
PT transplantation.  
XX  
PS Claim 10; Page 1125; 2508pp; English.  
XX  
CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins  
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-  
CC ABN79587 represent cDNAs encoding them. The invention also encompasses  
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively  
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to  
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX  
CC polynucleotides, the recombinant production of ORFX proteins, antibodies  
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
CC polypeptides, methods of screening for modulators of ORFX expression or  
CC activity, and methods of screening individuals for a predisposition to an  
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide  
CC range of biological activities, such as cytokine, cell proliferation,  
CC cell differentiation, immune modulation, haematopoiesis regulation,  
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/  
CC chemokinetic activity, haemostatic activity, thrombolytic activity,  
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
CC and antiinfective activity, and may also be involved in the determination  
CC of bodily characteristics, fertility and behaviour. ORFX proteins,  
CC nucleic acids and antibodies may be used in the treatment of cancers,  
CC other proliferative disorders such as psoriasis and benign tumours,  
CC neurological disorders such as epilepsy and Alzheimer's disease,  
CC cardiovascular diseases, immune system disorders, disorders related to  
CC organ transplantation, disorders of tissue growth and regeneration,  
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol  
CC storage disease, and infectious diseases caused by viral, bacterial,  
CC fungal and other pathogens. ORFX nucleic acids may also be used as a  
CC source of primers and probes, in the detection of ORFX genomic sequences  
CC or transcripts, in the identification and cloning of homologous  
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX  
CC nucleic acids may additionally be used to produce transgenic animals  
CC which may be useful for studying the function and/or activity of ORFX  
CC protein, and in drug screening. The ORFX proteins may also be used as  
CC immunogens to generate specific antibodies, which are useful in the  
CC diagnosis, treatment and monitoring of ORFX-associated diseases  
XX  
XX Sequence 112 AA;  
Query Match 92.9%; Score 26; DB 5; Length 112;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVEE 6  
DB 26 GYEVEE 31  
RESULT 46  
AAV28420  
ID AAV28420 standard; protein; 115 AA.  
XX  
XX AAV28420;  
XX  
XX 15-FEB-2000 (first entry)  
DE Corn branched chain amino acid transaminase amino acid sequence #3.  
XX  
XX Corn; soybean; wheat; rice; dihydroxyacid dehydratase; leuC; leuD;  
KW branched chain amino acid transferase; biosynthetic enzyme; antibody;  
KW 3-isopropylmalate dehydratase.  
XX  
XX Zea mays.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 104  
FT /label= unknown

FT Misc-difference 112 /note= "encoded by ANG"  
FT /label= unknown  
FT /note= "encoded by TAN"  
XX WO9921880-A2.  
XX  
XX 06-MAY-1999.  
XX  
XX 20-OCT-1998; 98WO-US022081.  
XX  
XX 28-OCT-1997; 97US-0063423P.  
XX  
XX (DUPO) DU PONT DE NEMOURS & CO E I.  
PI Falco SC, Cahoon RE, Hitz WD, Kinney AJ, Rafalski JA;  
XX  
XX WPI; 2000-022904/02.  
XX  
XX N-ESDB; AAX89447.  
XX  
XX Nucleic acid fragments encoding branched chain amino acid biosynthetic  
XX enzymes.  
XX  
XX Claim 6; Fig 3; 102pp; English.  
XX  
XX AAY28418-Y28431 are fragments of corn, soybean, wheat and rice branched  
XX chain amino acid transaminase, amino acid sequences. Sequences AAX89442-  
XX X89465 are nucleic acid fragments that encode all or a substantial  
XX portion of dihydroxyacid dehydratase, a branched chain amino acid  
XX transferase, a leuC or a leuD subunit of 3-isopropylmalate dehydratase  
XX from wheat, corn, soybean or rice. These enzymes are involved in  
XX biosynthesis and utilization of branched-chain amino acids. The nucleic  
XX acid sequences can be used to alter the level of expression of a branched  
XX chain amino acid biosynthetic enzyme in a host cell. They can also be  
XX used to obtain a nucleic acid fragment encoding a branched chain amino  
XX acid biosynthetic enzyme, and to identify inhibitors of a branched chain  
XX amino acid biosynthetic enzyme. The encoded proteins may be used to  
XX prepare antibodies for detecting the proteins in situ in cells, or in  
XX vitro in cell extracts  
XX  
XX Sequence 115 AA;  
Query Match 92.9%; Score 26; DB 3; Length 115;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVEE 6  
DB 33 GYRVEE 38  
RESULT 47  
ABB49265  
ID ABB49265 standard; protein; 116 AA.  
XX  
XX ABB49265;  
XX  
XX 05-FEB-2002 (first entry)  
XX  
XX Listeria monocytogenes protein #1969.  
XX  
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
KW vitamin B12; bacterial infection; disease.  
XX  
XX Listeria monocytogenes.  
XX  
XX WO200177335-A2.  
XX  
XX 18-OCT-2001.  
XX  
XX 11-APR-2001; 2001WO-FR001118.  
XX  
XX 11-APR-2000; 2000FR-00004629.  
XX  
XX

```

XX PA (INSP ) INST PASTEUR.
XX
XX PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsthi H, Dehoux P;
XX PI Duesurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
XX PI Daniels J, Goebel W, Krefte J, Kuhn M, Ng E, Vazquez-Boland JA;
XX PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
XX PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX PI Maduenio E, De Pablo B, Wehland J, Kaerst U, Entian K, Hauf J;
XX PI Rose M, Voss H;
XX
XX DR WPI; 2002-010914/01.
XX
XX PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX PT and prevention of Listeria and related bacterial infections, and related
XX PT polypeptides.
XX
XX PS Claim 6; SEQ ID NO 1970; 192pp; French.
XX
XX CC The present invention relates to the genome sequence of Listeria
XX CC monocytogenes EGP-e (see ABA03041). The genome sequence and fragments of
XX CC it are useful for selecting probes and primers for detecting genes in L.
XX CC monocytogenes and related organisms, and for studying genetic
XX CC polymorphisms and other genomes. The present sequence is a protein
XX CC encoded by the genome sequence of the present invention. Proteins
XX CC expressed from the genome sequence are useful for raising specific
XX CC antibodies, identification of L. monocytogenes and related organisms, and
XX CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX CC B12. The genome sequence and proteins encoded by it are also useful for
XX CC selecting compounds that regulate gene expression and cell replication
XX CC and modulate L. monocytogenes-related diseases. In addition, the genome
XX CC sequence and proteins encoded by it are useful in pharmaceutical and
XX CC vaccines compositions for the treatment or prevention of infections by L.
XX CC monocytogenes and related organisms. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 116 AA;
XX
XX Query Match 92.9%; Score 26; DB 5; Length 116;
XX Best Local Similarity 83.3%; Pred. No. 1.5e+02;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GYXVEE 6
XX Db |||||
XX 80 GYRVEE 85
XX
XX RESULT 48
XX AAB48256
XX ID AAB48256 standard; protein; 118 AA.
XX
XX AC AAB48256;
XX
XX DT 02-APR-2001 (first entry)
XX
XX DE Corn magnesium chelataze subunit (clone cdtic.pk001.ol).
XX
XX KW Magnesium chelataze; transgenic; herbicide; gene marker; plant breeding;
XX KW corn.
XX
XX OS Zea mays.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 14
XX FT /note= "the corresponding nucleotide sequence encodes a
XX FT residue pro at this position which is not indicated in
XX FT this sequence"
XX FT Misc-difference 101
XX FT /note= "encoded by GC"
XX FT
XX

```

```

PN WO200075340-A2.
XX
XX PD 14-DEC-2000.
XX
XX PF 02-JUN-2000; 2000WO-US015351.
XX
XX PR 04-JUN-1999; 99US-0137461P.
XX
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX PI Butler KH, Famodu OO, Gutteridge S, Maxwell CA;
XX
XX DR WPI; 2001-091215/10.
XX DR N-FSDB; AAC84575.
XX
XX PT Isolated nucleic acid fragments encoding magnesium chelataze subunits,
XX PT useful as probes for genetic and physical mapping of genes, as markers
XX PT for traits linked to these genes, and in plant breeding.
XX
XX PS Claim 10; Page 67-68; 103pp; English.
XX
XX CC The invention relates to nucleic acid fragments encoding magnesium
XX CC chelataze subunits. The nucleic acid fragments may be used to create
XX CC transgenic plants in which the new polypeptides are present at higher or
XX CC lower levels than normal or in cell types or developmental stages in
XX CC which they are not normally found, and for overexpression in bacterial or
XX CC yeast hosts to efficiently produce large amounts of the encoded
XX CC polypeptides which could then be used for screening different compounds
XX CC for potential herbicidal activity. The polynucleotides may also be used
XX CC as probes for genetic and physical mapping the genes that they are part
XX CC of, and as markers for traits linked to these genes. Such information is
XX CC useful in plant breeding. The polypeptides are used for preparing
XX CC antibodies, which are useful for detecting the polypeptides in situ or in
XX CC vitro, and as a target to facilitate design and/or identification of
XX CC inhibitors of enzymes that may be used as herbicides. Host cells may also
XX CC be used directly for screening different compounds for potential
XX CC herbicidal activity. The present sequence represents a corn magnesium
XX CC chelataze subunit
XX
XX SQ Sequence 118 AA;
XX
XX Query Match 92.9%; Score 26; DB 4; Length 118;
XX Best Local Similarity 83.3%; Pred. No. 1.5e+02;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GYXVEE 6
XX Db |||||
XX 23 GYLVEE 28
XX
XX RESULT 49
XX ADH85685
XX ID ADH85685 standard; protein; 128 AA.
XX
XX AC ADH85685;
XX
XX DT 22-APR-2004 (first entry)
XX
XX DE Enterococcus faecalis polypeptide #165.
XX
XX KW Enterococcus faecalis infection; transcription regulatory element;
XX KW antibacterial.
XX
XX OS Enterococcus faecalis.
XX
XX PN US6617156-B1.
XX
XX PD 09-SEP-2003.
XX
XX PF 13-AUG-1998; 98US-00134000.
XX
XX PR 15-AUG-1997; 97US-0055778P.
XX

```

PA (DOUC/) DOUCETTE-STAMM L A.  
 PA (BUSH/) BUSH D.  
 FI Doucette-Stamm LA, Bush D;  
 XX WPI; 2003-895394/82.  
 XX N-PSDB; ADH82280.  
 PT New nucleic acid comprising a sequence encoding an Enterococcus faecalis  
 PT polypeptide, useful for preparing a composition for diagnosing or  
 PT treating E. faecalis infection.  
 XX Disclosure; SEQ ID NO 3570; 193pp; English.  
 XX The invention relates to Enterococcus faecalis polynucleotides and  
 CC polypeptides. The invention also relates to a recombinant expression  
 CC vector comprising a polynucleotide operably linked to a transcription  
 CC regulatory element, a cell comprising a recombinant vector, a method for  
 CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising  
 CC a sequence not given in the specification, a recombinant vector  
 CC comprising the nucleic acid and a cell comprising the recombinant vector.  
 CC The polynucleotides can be used to detect the presence of E. faecalis in  
 CC a sample. The sequences are useful for preparing a composition for  
 CC diagnosing or treating Enterococcus faecalis infection. This sequence  
 CC represents an E. faecalis polypeptide of the invention.  
 XX SQ Sequence 128 AA;  
 Query Match 92.9%; Score 26; DB 7; Length 128;  
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVEE 6  
 DB 92 GYRVEE 97  
 RESULT 50  
 AAM15152  
 ID AAM15152 standard; protein; 139 AA.  
 AC AAM15152;  
 XX 12-OCT-2001 (first entry)  
 DE Peptide #1586 encoded by probe for measuring cervical gene expression.  
 XX Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX Homo sapiens.  
 XX WO200157278-A2.  
 PN 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US000670.  
 PF 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488901/53.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human cervical epithelial cells.  
 XX Claim 27; SEQ ID NO 19978; 487pp; English.  
 XX The present invention relates to human single exon nucleic acid probes  
 CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 139 AA;  
 Query Match 92.9%; Score 26; DB 4; Length 139;  
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVEE 6  
 DB 11 GYDVEE 16  
 RESULT 51  
 AAM27609  
 ID AAM27609 standard; protein; 139 AA.  
 XX AAM27609;  
 AC AAM27609;  
 XX 17-OCT-2001 (first entry)  
 DE Peptide #1646 encoded by probe for measuring placental gene expression.  
 XX Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX Homo sapiens.  
 XX WO200157272-A2.  
 PN 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US000663.  
 PF 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-48897/53.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.  
 XX Claim 27; SEQ ID NO 27878; 654pp; English.  
 XX The present invention relates to single exon nucleic acid probes (SENP:  
 CC see AAI1315-AAI57546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders



```
SQ Sequence 139 AA;
Query Match          92.9%; Score 26; DB 4; Length 139;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
   |||||
DB 11 GYDVEE 16

RESULT 52
ABB28977
ID ABB28977 standard; peptide; 139 AA.
XX
AC ABB28977;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #1628 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 27; SEQ ID NO 11945; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast. And then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 139 AA;
Query Match          92.9%; Score 26; DB 4; Length 139;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
   |||||
DB 11 GYDVEE 16

RESULT 53
ABB19587
ID ABB19587 standard; protein; 139 AA.
XX
AC ABB19587;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #1586 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 21357; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC AB221535-ABA1305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease.
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 139 AA;
Query Match          92.9%; Score 26; DB 4; Length 139;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
   |||||
DB 11 GYDVEE 16
```

```

RESULT 54
AAM67316
ID AAM67316 standard; protein; 139 AA.
XX AC AAM67316;
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27622.
XX DE Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488900/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human bone marrow.
XX PS Example 4; SEQ ID NO 27622; 658pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a
XX CC protein encoded by one of the probes of the invention
XX SQ Sequence 139 AA;

Query Match 92.9%; Score 26; DB 4; Length 139;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 11 GYDVEE 16

RESULT 55
AAM54935
ID AAM54935 standard; protein; 139 AA.
XX AC AAM54935;
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27040.
XX DE Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX OS Homo sapiens.

```

```

XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483446/52.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains.
XX PS Example 4; SEQ ID NO 27040; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, of
XX CC epilepsy and cancers. The present sequence is a protein encoded by one of
XX CC the probes of the invention
XX SQ Sequence 139 AA;

Query Match 92.9%; Score 26; DB 4; Length 139;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 11 GYDVEE 16

RESULT 56
ABG48978
ID ABG48978 standard; peptide; 139 AA.
XX AC ABG48978;
XX DT 25-FEB-2003 (first entry)
XX DE Human liver peptide, SEQ ID NO 27626.
XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX KW hypercholesterolaemia; coronary heart disease.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.

```

XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-488898/53.  
XX XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.  
XX PS Claim 27; SEQ ID NO 27626; 658pp; English.  
XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridizes at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (I) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABG47348-ABG59930 represent human  
CC liver single exon encoded peptides of the invention. Note: The sequence  
CC information for this patent does not appear in the printed specification  
CC but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 139 AA;  
Query Match 92.9%; Score 26; DB 4; Length 139;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GYXVEE 6  
Db 11 GYDVEE 16  
|||  
|||  
RESULT 57  
ABO57215  
ID ABO57215 standard; protein; 142 AA.  
XX AC ABO57215;  
XX DT 29-JUL-2004 (first entry)  
XX DE Human genome derived single exon protein #3449.  
XX KW Human; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.  
XX OS Homo sapiens.  
XX PN US2003194704-A1.  
XX XX 16-OCT-2003.  
XX PF 03-APR-2002; 2002US-00029386.  
XX PR 03-APR-2002; 2002US-00029386.  
XX XX (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX PI Penn SG, Rank DR, Hanzel DK;  
XX DR WPI; 2004-119264/12.  
XX PT New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.

XX PS Claim 45; SEQ ID NO 30849; 80pp; English.  
XX CC The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above). The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe protein of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20030194704  
XX SQ Sequence 142 AA;  
Query Match 92.9%; Score 26; DB 8; Length 142;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GYXVEE 6  
Db 56 GYLVEE 61  
|||  
|||  
RESULT 58  
AAB57021  
ID AAB57021 standard; protein; 154 AA.  
XX AC AAB57021;  
XX DT 13-MAR-2001 (first entry)  
XX DE Human prostate cancer antigen protein sequence SEQ ID NO:1599.  
XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
KW wound; infectious disease.  
XX OS Homo sapiens.  
XX PN WO200055174-A1.  
XX XX 21-SEP-2000.  
XX PT



XX 05-FEB-2002 (first entry)  
 XX Listeria monocytogenes protein #546.  
 DE Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
 KW vitamin B12; bacterial infection; disease.  
 XX Listeria monocytogenes.  
 OS WO200177335-A2.  
 XX 18-OCT-2001.  
 PD 11-APR-2001; 2001WO-FR001118.  
 XX 11-APR-2000; 2000FR-00004629.  
 PR (INSP ) INST PASTEUR.  
 PA Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;  
 PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;  
 PI Daniels J, Goebel W, Kretz J, Kuhn M, Ng E, Vazquez-Boland JA;  
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;  
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
 PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;  
 PI Rose M, Voss H;  
 DR WPI; 2002-010914/01.  
 XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment  
 PT and prevention of *Listeria* and related bacterial infections, and related  
 PT polypeptides.  
 XX Claim 6; SEQ ID NO 547; 192pp; French.  
 XX The present invention relates to the genome sequence of *Listeria*  
 CC *monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of  
 CC it are useful for selecting probes and primers for detecting genes in *L.*  
 CC *monocytogenes* and related organisms, and for studying genetic  
 CC polymorphisms and other genomes. The present sequence is a protein  
 CC encoded by the genome sequence of the present invention. Proteins  
 CC expressed from the genome sequence are useful for raising specific  
 CC antibodies, identification of *L. monocytogenes* and related organisms, and  
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin  
 CC B12. The genome sequence and proteins encoded by it are also useful for  
 CC selecting compounds that regulate gene expression and cell replication  
 CC and modulate *L. monocytogenes*-related diseases. In addition, the genome  
 CC sequence and proteins encoded by it are useful in pharmaceutical and  
 CC vaccine compositions for the treatment or prevention of infections by *L.*  
 CC *monocytogenes* and related organisms. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 161 AA;  
 SQ Query Match 92.9%; Score 26; DB 5; Length 161;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 GVXVEE 6  
 DB 125 GYATIE 130  
 RESULT 61  
 ID ABO59796  
 XX ABO59796 standard; protein; 173 AA.  
 AC ABO59796;  
 XX

DT 29-JUL-2004 (first entry)  
 XX Human genome derived single exon protein #6030.  
 DE Human; gene expression; single exon probe; microarray;  
 KW alternative splicing event; genomic alteration.  
 XX Homo sapiens.  
 OS US2003194704-A1.  
 XX 16-OCT-2003.  
 PD 03-APR-2002; 2002US-00029386.  
 PF 03-APR-2002; 2002US-00029386.  
 PR (PENN/) PENN S G.  
 PA (RANK/) RANK D R.  
 PA (HANK/) HANZEL D K.  
 PI Penn SG, Rank DR, Hanzel DK;  
 DR WPI; 2004-119264/12.  
 XX New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for  
 PT surveying tissues.  
 XX Claim 45; SEQ ID NO 33430; 80pp; English.  
 XX The invention relates to a nucleic acid probe for measuring human gene  
 CC expression, comprising any of the 27,400 fully defined nucleotide  
 CC sequences in the specification, or their complements or fragments, and  
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
 CC fully defined in the specification. The probe is a single exon probe that  
 CC hybridises under high stringency conditions to a nucleic acid molecule  
 CC expressed in human cells or tissues. Also included are a spatially-  
 CC addressable set of single exon nucleic acid probes for measuring human  
 CC gene expression (comprising a plurality of single exon nucleic acid  
 CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable from the plurality), a single  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC methods of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing  
 CC human gene expression data by subscription, and a computer-readable  
 CC storage medium which contains a database having a plurality of records  
 CC (each record including data on the expression of a single exon probe  
 CC cited above). The probe, methods and apparatus are useful in gene  
 CC expression analysis. The probes may be used as tools for surveying  
 CC tissues to detect the presence of expressed messages that contain their  
 CC specific exon, or in constructing genome-derived single exon microarrays.  
 CC In addition, the probes are used in identifying and characterising  
 CC alternative splicing events, in detecting and characterising gross  
 CC alterations in the genomic locus that includes their exon, in assessing  
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
 CC or in expressing the ORF-encoded peptide. The present sequence is a human  
 CC single exon probe protein of the invention. Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docID=20030194704  
 XX Sequence 173 AA;  
 SQ Query Match 92.9%; Score 26; DB 8; Length 173;  
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy	1	GYXVEE 6	
Db	77	GYLVEE 82	
RESULT 62			
ABGI7723			
ID	ABGI7723	standard; protein; 188 AA.	
XX	AC	ABGI7723;	
XX	AC	ABGI7723;	
XX	DT	18-FEB-2002 (first entry)	
XX	DE	Novel human diagnostic protein #17714.	
XX	DE	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
XX	KW	food supplement; medical imaging; diagnostic; genetic disorder.	
XX	KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
XX	KW	food supplement; medical imaging; diagnostic; genetic disorder.	
XX	OS	Homo sapiens.	
XX	OS	WO200175067-A2.	
XX	PN	11-OCT-2001.	
XX	PD	30-MAR-2001; 2001WO-US008631.	
XX	PF	31-MAR-2000; 2000US-00540217.	
XX	PR	23-AUG-2000; 2000US-00649167.	
XX	PR	(HYSE-) HYSEQ INC.	
XX	PA	Drmanac RT, Liu C, Tang YT;	
XX	PI	WPI; 2001-639362/73.	
XX	DR	N-PSDB; AAS81910.	
XX	XX	New isolated polynucleotide and encoded polypeptides, useful in	
PT	PT	diagnostics, forensics, gene mapping, identification of mutations	
PT	PT	responsible for genetic disorders or other traits and to assess	
PT	PT	biodiversity.	
XX	XX	Claim 20; SEQ ID NO 48082; 103pp; English.	
XX	XX	The invention relates to isolated polynucleotide (I) and polypeptide (II)	
CC	CC	sequences. (II) is useful as hybridisation probes, polymerase chain	
CC	CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,	
CC	CC	and in recombinant production of (II). The polynucleotides are also used	
CC	CC	in diagnostics as expressed sequence tags for identifying expressed	
CC	CC	genes. (I) is useful in gene therapy techniques to restore normal	
CC	CC	activity of (II) or to treat disease states involving (II). (II) is	
CC	CC	useful for generating antibodies against it, detecting or quantitating a	
CC	CC	polypeptide in tissue, as molecular weight markers and as a food	
CC	CC	supplement. (II) and its binding partners are useful in medical imaging	
CC	CC	of sites expressing (II). (I) and (II) are useful for treating disorders	
CC	CC	involving aberrant protein expression or biological activity. The	
CC	CC	polypeptide and polynucleotide sequences have applications in	
CC	CC	diagnostics, forensics, gene mapping, identification of mutations	
CC	CC	responsible for genetic disorders or other traits to assess biodiversity	
CC	CC	and to produce other types of data and products dependent on DNA and	
CC	CC	amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic	
CC	CC	amino acid sequences of the invention. Note: The sequence data for this	
CC	CC	patent did not appear in the printed specification, but was obtained in	
CC	CC	electronic format directly from WIPO at	
CC	CC	ftp.wipo.int/pub/published_pct_sequences	
XX	XX	Sequence 188 AA;	
SQ			
Query Match	92.9%;	Score 26; DB 4; Length 188;	
Best Local Similarity	83.3%;	Pred. No. 2.6e+02;	
Matches	5; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
Oy	1	GYXVEE 6	

PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.

PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156536P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 06-OCT-1999; 99US-015753P.  
PR 07-OCT-1999; 99US-0157865P.  
PR 08-OCT-1999; 99US-0158029P.  
PR 12-OCT-1999; 99US-0158232P.  
PR 13-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160788P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.9%; Score 26; DB 3; Length 194;  
Best Local Similarity 83.3%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYVEE 6  
:  
DB 181 GYVEE 186

RESULT 64  
ADG27817  
ID ADG27817 standard; protein; 197 AA.

XX AC ADG27817;  
XX DT 26-FEB-2004 (first entry)  
XX DE Human novel protein amino acid sequence SeqID584.  
XX KW antiparkinsonian; haemostatic; nootropic; neuroprotective; osteopathic;  
KW anti-HIV; protozoacide; antifungal; immunosuppressive; antirheumatic;  
KW antiarthritic; antidiabetic; antiallergic; antiinflammatory;  
KW anticoagulant; cytostatic; gene therapy; Parkinson's disease;  
KW Alzheimer's disease; thrombocytopaenia; osteoporosis; osteoarthritis;  
KW infection, HIV; Leishmania; malaria; fungal infection;  
KW multiple sclerosis; rheumatoid arthritis;  
KW insulin dependent diabetes mellitus; allergic reaction; food allergy;  
KW insect allergy; allergic rhinitis; haemophilia; cancer; human.  
XX OS Homo sapiens.  
XX WO200179254-A1.  
XX PD 25-OCT-2001.  
XX PF 16-APR-2001; 2001WO-US008655.  
XX PR 18-APR-2000; 2000US-00552929.  
XX PR 22-SEP-2000; 2000US-00668317.  
XX PR 24-OCT-2000; 2000US-00695783.  
XX PR 01-DEC-2000; 2000US-00728628.  
XX PR 26-JAN-2001; 2001US-00770160.  
XX PR 13-FEB-2001; 2001US-00783066.  
XX PR 22-MAR-2001; 2001US-00816828.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;  
PI Yang Y, Zhao QA, Goodrich RW, Liu C, Drmanac RT, Ma Y, Wang Z;  
PI Wehrman T;  
XX WPI; 2001-607699/69.  
XX Novel polynucleotides and encoded polypeptides (protein factors,  
PT including, e.g., cytokines, such as lymphokines, interferons, and  
PT circulating soluble factors) useful for treating, e.g., Parkinson's,  
PT Alzheimer's, HIV and cancer.  
XX Claim 20; SEQ ID NO 584; 153pp; English.  
XX This invention relates to a novel isolated DNA sequence and the mature  
CC proteins encoded by them. The invention may be useful in the development  
CC of compositions with antiparkinsonian, haemostatic, nootropic,  
CC neuroprotective, osteopathic, anti-HIV, protozoacide, antifungal,  
CC immunosuppressive, antirheumatic, antiarthritic, antidiabetic,  
CC antiallergic, antiinflammatory, anticoagulant or cytostatic activities.  
CC In addition, the sequences of the invention may be useful for gene  
CC therapy. The invention may be useful for the development of treatments  
CC for Parkinson's, Alzheimer's, thrombocytopaenia, osteoporosis,  
CC osteoarthritis, infections (including HIV, Leishmania, malaria, and  
CC various fungal infection), autoimmune disorders such as multiple  
CC sclerosis, rheumatoid arthritis, and insulin dependent diabetes mellitus,  
CC allergic reactions and conditions (for example food allergies, insect  
CC allergies and allergic rhinitis), coagulation disorders including  
CC haemophilia, and cancer. Note: The amino acid sequences given in table 6  
CC (SeqID 439-584) may have in frame stop codons or possible  
CC insertions/deletions as shown in the table. The sequences allocated Seq  
CC IDs 1-438 are not provided (even by reference) in the specification.  
XX Sequence 197 AA;

Query Match 92.9%; Score 26; DB 4; Length 197;  
Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
XX ||| |||  
Db 181 GYLVEE 186  
RESULT 65  
AAG43322  
ID AAG43322 standard; protein; 199 AA.  
XX AAG43322;  
AC AAG43322;  
DT 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 54135.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
XX PR 05-MAR-1999; 99US-0123180P.  
XX PR 09-MAR-1999; 99US-0123548P.  
XX PR 23-MAR-1999; 99US-0125788P.  
XX PR 25-MAR-1999; 99US-0126264P.  
XX PR 29-MAR-1999; 99US-0126785P.  
XX PR 01-APR-1999; 99US-0127462P.  
XX PR 06-APR-1999; 99US-0128234P.  
XX PR 08-APR-1999; 99US-0128714P.  
XX PR 16-APR-1999; 99US-0129845P.  
XX PR 21-APR-1999; 99US-0130077P.  
XX PR 23-APR-1999; 99US-0130510P.  
XX PR 28-APR-1999; 99US-0130891P.  
XX PR 30-APR-1999; 99US-0132048P.  
XX PR 04-MAY-1999; 99US-0132407P.  
XX PR 05-MAY-1999; 99US-0132484P.  
XX PR 06-MAY-1999; 99US-0132486P.  
XX PR 07-MAY-1999; 99US-0132487P.  
XX PR 11-MAY-1999; 99US-0132863P.  
XX PR 14-MAY-1999; 99US-0134218P.  
XX PR 14-MAY-1999; 99US-0134219P.  
XX PR 14-MAY-1999; 99US-0134221P.  
XX PR 18-MAY-1999; 99US-0134370P.  
XX PR 19-MAY-1999; 99US-0134941P.  
XX PR 20-MAY-1999; 99US-0135124P.  
XX PR 21-MAY-1999; 99US-0135353P.  
XX PR 25-MAY-1999; 99US-0135629P.  
XX PR 27-MAY-1999; 99US-0136021P.  
XX PR 28-MAY-1999; 99US-0136782P.  
XX PR 01-JUN-1999; 99US-0137222P.  
XX PR 03-JUN-1999; 99US-0137528P.  
XX PR 04-JUN-1999; 99US-0137502P.  
XX PR 07-JUN-1999; 99US-0137724P.  
XX PR 08-JUN-1999; 99US-0138094P.  
XX PR 10-JUN-1999; 99US-0138540P.  
XX PR 10-JUN-1999; 99US-0138847P.  
XX PR 14-JUN-1999; 99US-0139119P.  
XX PR 16-JUN-1999; 99US-0139452P.  
XX PR 16-JUN-1999; 99US-0139453P.  
XX PR 17-JUN-1999; 99US-0139492P.



PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145911P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.

PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149417P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.9%; Score 26; DB 3; Length 199;  
Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
Db 186 GYVEE 191

RESULT 66  
ABG02504

ID ABG02504 standard; protein; 213 AA.  
XX AC  
XX AC ABG02504;  
XX DT  
XX DT 13-FEB-2002 (first entry)  
XX DE  
XX DE Novel human diagnostic protein #2495.  
XX XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX KW  
XX OS Homo sapiens.  
XX XX  
XX PN WO200175067-A2.  
XX PD  
XX PD 11-OCT-2001.  
XX PF  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX PR  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX XX  
XX PA (HYSE-) HYSEQ INC.  
XX PI  
XX PI Drmanac RT, Liu C, Tang YT;  
XX XX  
XX DR WPI; 2001-639362/73.  
XX DR N-PSDB; AAS66691.  
XX XX  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity.  
XX PS  
XX PS Claim 20; SEQ ID NO 32863; 103pp; English.  
XX XX  
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain  
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX CC and in recombinant production of (II). The polynucleotides are also used  
XX CC in diagnostics as expressed sequence tags for identifying expressed  
XX CC genes. (I) is useful in gene therapy techniques to restore normal  
XX CC activity of (II) or to treat disease states involving (II). (II) is  
XX CC useful for generating antibodies against it, detecting or quantitating a  
XX CC polypeptide in tissue, as molecular weight markers and as a food  
XX CC supplement. (II) and its binding partners are useful in medical imaging  
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders  
XX CC involving aberrant protein expression or biological activity. The  
XX CC polypeptide and polynucleotide sequences have applications in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC and to produce other types of data and products dependent on DNA and  
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
XX CC amino acid sequences of the invention. Note: The sequence data for this  
XX CC patent did not appear in the printed specification, but was obtained in  
XX CC electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX XX  
XX SQ Sequence 213 AA;  
  
Query Match 92.9%; Score 26; DB 4; Length 213;  
Best Local Similarity 83.3%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVEE 6  
DB 77 GYLVEE 82  
  
RESULT 67  
AAV13620  
ID AAY13620 standard; protein; 218 AA.  
XX XX

AC AAY13620;  
XX DT  
XX DT 03-SEP-1999 (first entry)  
XX DE  
XX DE Vegetative storage protein variant VSPbeta-Met10 sequence.  
XX KW  
XX KW Engineered protein; nutritional value; plant protein; VSP;  
XX KW vegetative storage protein; variant.  
XX OS  
XX OS Glycine max.  
XX OS Synthetic.  
XX PN  
XX PN WO9929882-A2.  
XX PD  
XX PD 17-JUN-1999.  
XX PF  
XX PF 10-DEC-1998; 98WO-US026209.  
XX PR  
XX PR 10-DEC-1997; 97US-00988015.  
XX PA (PTON-) PIONEER HI-BRED INT INC.  
XX PI  
XX PI Rao GA, Sleister HM;  
XX DR  
XX DR WPI; 1999-385617/32.  
XX DR N-PSDB; AAX55798.  
XX PT  
XX PT Altering the amino acid composition of proteins, particularly in plants.  
XX PS  
XX PS Disclosure; Fig 2; 42pp; English.  
XX CC  
XX CC The invention describes a new method for altering the amino acid  
XX CC composition of proteins, particularly in plants. The method comprises  
XX CC introducing changes to create an engineered protein (EP) having the same  
XX CC conformation as native protein, and can be used for increasing  
XX CC nutritional value. The methods can be used to increase in EPS the levels  
XX CC of essential amino acids e.g. methionine, tryptophan, lysine, valine,  
XX CC phenylalanine, isoleucine, threonine or cysteine. They can be used for  
XX CC increasing the nutritional value of plant proteins. They can be used in  
XX CC dicot and monocot plants, e.g. maize or soybean. Sequences AAY13620-622  
XX CC represent proposed methionine-enriched variants of a vegetative storage  
XX CC protein (VSP) VSPbeta  
XX SQ Sequence 218 AA;  
  
Query Match 92.9%; Score 26; DB 2; Length 218;  
Best Local Similarity 83.3%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVEE 6  
DB 87 GYVEE 92  
  
RESULT 68  
AAV13619  
ID AAY13619 standard; protein; 218 AA.  
XX AC  
XX AC AAY13619;  
XX DT  
XX DT 03-SEP-1999 (first entry)  
XX DE  
XX DE Vegetative storage protein VSPbeta sequence.  
XX KW  
XX KW Engineered protein; nutritional value; plant protein; VSP;  
XX KW vegetative storage protein.  
XX OS  
XX OS Glycine max.  
XX PN  
XX PN WO9929882-A2.  
XX PD  
XX PD 17-JUN-1999.

PF 10-DEC-1998; 98WO-US026209.  
 XX 10-DEC-1997; 97US-00988015.  
 PR (PION-) PIONEER HI-BRED INT INC.  
 XX Rao GA, Sleister HM;  
 PI WPI; 1999-385617/32.  
 DR Altering the amino acid composition of proteins, particularly in plants.  
 XX Disclosure; Fig 2; 42pp; English.  
 PS The invention describes a new method for altering the amino acid  
 CC composition of proteins, particularly in plants. The method comprises  
 CC introducing changes to create an engineered protein (EP) having the same  
 CC conformation as native protein, and can be used for increasing  
 CC nutritional value. The methods can be used to increase in Epe the levels  
 CC of essential amino acids e.g. methionine, tryptophan, lysine, valine,  
 CC phenylalanine, isoleucine, threonine or cysteine. They can be used for  
 CC increasing the nutritional value of plant proteins. They can be used for  
 CC dicot and monocot plants, e.g. maize or soybean. The present sequence  
 CC represents a vegetative storage protein (VSP) VSPbeta  
 XX Sequence 218 AA;  
 SQ

Query Match 92.9%; Score 26; DB 2; Length 218;  
 Best Local Similarity 83.3%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 |||||  
 Db 87 GYGVVE 92

RESULT 69  
 ABG02618  
 ID ABG02618 standard; protein; 218 AA.  
 XX  
 AC ABG02618;  
 XX  
 DT 13-FEB-2002 (first entry)  
 DE Novel human diagnostic protein #2609.  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX Homo sapiens.  
 OS WO200175067-A2.  
 PN 11-OCT-2001.  
 XX  
 PD 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR N-PSDB; AAS66805.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PS Claim 20; SEQ ID NO 32977; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 XX ftp.wipo.int/pub/published\_pct\_sequences  
 SQ Sequence 218 AA;  
 CC

Query Match 92.9%; Score 26; DB 4; Length 218;  
 Best Local Similarity 83.3%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 |||||  
 Db 70 GYLVEE 75

RESULT 70  
 ABG02492  
 ID ABG02492 standard; protein; 218 AA.  
 XX  
 AC ABG02492;  
 XX  
 DT 13-FEB-2002 (first entry)  
 DE Novel human diagnostic protein #2483.  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX Homo sapiens.  
 OS WO200175067-A2.  
 PN 11-OCT-2001.  
 XX  
 PD 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR N-PSDB; AAS66679.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PS Claim 20; SEQ ID NO 32851; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC

sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (III) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activities in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 218 AA;

Query Match 92.9%; Score 26; DB 4; Length 218;  
Best Local Similarity 83.3%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 70 GYLVEE 75

RESULT 71  
ADN46899  
ID ADN46899 standard; protein; 231 AA.  
XX  
AC ADN46899;  
DT 01-JUL-2004 (first entry)  
XX Thermococcus kodakaraensis KOD1 protein sequence SeqID777.  
DE  
XX gene disruption; gene targeting; marker gene; transformation;  
KW homologous recombination; hyperthermostable archaeobacterium; KOD1;  
KW gene structure; gene function; enzyme activity; medicine;  
KW forensic science; food; drug inspection; molecular biology; immunology.  
XX  
OS Thermococcus kodakaraensis.  
XX WO2004022736-A1.  
PN  
XX 18-MAR-2004.  
PD  
XX 29-AUG-2003; 2003WO-IB003597.  
PF  
XX 30-AUG-2002; 2002JP-00319011.  
PR  
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
PA  
XX Imanaka T, Atomi H;  
PI  
XX WPI; 2004-257583/24.  
DR  
XX Method for disrupting targeted gene in genome of organism particularly  
PT thermotable bacterium and with genome chips for analysis, applicable in  
PT studying gene structure and functions.  
XX  
PS Claim 9; SEQ ID NO 777; 598bp; Japanese.  
XX  
XX This invention relates to a novel method for targeting disruption of an  
CC arbitrary gene in a genome of an organism which comprises providing the  
CC whole sequential data of the genome of such organism, selecting at least  
CC 1 arbitrary region in the sequence, providing a vector that contains a

sequence homologous with the selected region and a marker gene, transformation, and homologous recombination. The genome is preferably the genome of a hyperthermostable archaeobacterium, particularly Thermococcus kodakaraensis KOD1. The method is for targeting the disruption of a gene in the genome of an organism, which is applicable in studying gene structure and functions as well as enzyme activities of encoded proteins and useful in medicine, forensic science, food or drug inspection, molecular biology and immunology. With this method, the disruption of a gene at an arbitrary position in a genome can be achieved efficiently and reliably. The present sequence is that of a protein encoded by the genome of Thermococcus kodakaraensis which was derived using the method of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 231 AA;

Query Match 92.9%; Score 26; DB 8; Length 231;  
Best Local Similarity 66.7%; Pred. No. 3.3e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 18 GYSIEE 23

RESULT 72  
ABB64001  
ID ABB64001 standard; protein; 233 AA.  
XX  
AC ABB64001;  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 18795.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL08104.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 18795; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 233 AA;

Query Match 92.9%; Score 26; DB 4; Length 233;  
Best Local Similarity 66.7%; Pred. No. 3.3e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 198 GYSIEE 203

RESULT 73  
ADC96641  
ID ADC96641 standard; protein; 238 AA.

XX AC ADC96641;

XX 01-JAN-2004 (first entry)

XX E. faecium protein sequence SEQ ID 6268.

XX Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;  
KW abdominal-pelvic infection.

XX Enterococcus faecium.

XX US6583275-B1.

XX 24-JUN-2003.

XX 30-JUN-1998; 98US-00107532.

XX 02-JUL-1997; 97US-0051571P.

XX 14-MAY-1998; 98US-0085598P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2003-799836/75.

XX N-PSDB; ADC92987.

PT New isolated nucleic acid derived from Enterococcus faecium encoding an  
PT Enterococcus faecium polypeptide useful for detection, prevention and  
PT treatment of a pathological condition resulting from a bacterial  
PT infection.

XX Example 1; SEQ ID NO 6268; 243pp; English.

XX The invention relates to an isolated nucleic acid derived from  
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having  
CC one of 10 fully defined sequences given in the (or comprising 40  
CC sequential nucleotides chosen from any of the nucleic acids, its  
CC complement or sequences hybridising to it). Also included are a  
CC recombinant vector comprising the nucleic acid operably linked to  
CC transcription regulatory element, a cell comprising the vector and a  
CC single-stranded probe comprising the nucleic acid. The nucleic acids are  
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.  
CC The nucleic acids are useful for diagnosing pathological conditions  
CC resulting from E. faecium bacterial infection (e.g. urinary tract  
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic  
CC infection) and for screening drugs such as agonists and antagonists. The  
CC nucleic acid is useful for recombinant production of Candida albicans -  
CC derived peptides or antisense polypeptides. Pharmaceutical compositions  
CC and vaccines containing the nucleic acid are useful for preventing or  
CC treating Enterococcus faecium infections. The present sequence represents  
CC one if the disclosed E. faecium proteins.

XX Sequence 238 AA;

Query Match 92.9%; Score 26; DB 7; Length 238;

Best Local Similarity 83.3%; Pred. No. 3.4e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 168 GYVEE 173

RESULT 74  
ABG08795  
ID ABG08795 standard; protein; 239 AA.

XX AC ABG08795;

XX 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #8786.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS72982.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 20; SEQ ID NO 39154; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 239 AA;

Query Match 92.9%; Score 26; DB 4; Length 239;  
Best Local Similarity 83.3%; Pred. No. 3.4e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
Db 102 GYLVEE 107

RESULT 75  
ABB49186  
ID ABB49186 standard; protein; 242 AA.  
XX  
AC ABB49186;  
XX  
DT 05-FEB-2002 (first entry)  
XX  
DE Listeria monocytogenes protein #1890.  
XX  
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
KW vitamin B12; bacterial infection; disease.  
XX  
OS Listeria monocytogenes.  
XX  
PN WO200177335-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 11-APR-2001; 2001WO-FR001118.  
XX  
PR 11-APR-2000; 2000FR-00004629.  
XX  
PA (INSP ) INST PASTEUR.  
XX  
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;  
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst P, Cossart P;  
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;  
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
PI Chakraborty T, Doman E, Hain T, Berche P, Charbit A, Durant L;  
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;  
PI Rose M, Voss H;  
XX  
WPI; 2002-010914/01.  
XX  
Genomic sequence for Listeria monocytogenes, useful e.g. for treatment  
PT and prevention of Listeria and related bacterial infections, and related  
PT polypeptides.  
XX  
PS Claim 6; SEQ ID NO 1891; 192pp; French.  
XX  
The present invention relates to the genome sequence of Listeria  
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of  
CC it are useful for selecting probes and primers for detecting genes in L.  
CC monocytogenes and related organisms, and for studying genetic  
CC polymorphisms and other genomes. The present sequence is a protein  
CC encoded by the genome sequence of the present invention. Proteins  
CC expressed from the genome sequence are useful for raising specific  
CC antibodies, identification of L. monocytogenes and related organisms, and  
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin  
CC B12. The genome sequence and proteins encoded by it are also useful for  
CC selecting compounds that regulate gene expression and cell replication  
CC and modulate L. monocytogenes-related diseases. In addition, the genome  
CC sequence and proteins encoded by it are useful in pharmaceutical and  
CC vaccines compositions for the treatment or prevention of infections by L.  
CC monocytogenes and related organisms. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 242 AA;

Qy 1 GYXVEE 6  
Db 214 GYIVEE 219

Search completed: November 1, 2004, 21:30:15  
Job time : 168 secs

Query Match 92.9%; Score 26; DB 5; Length 242;  
Best Local Similarity 83.3%; Pred. No. 3.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:15:40 ; Search time 24 Seconds  
(without alignments)  
24.054 Million cell updates/sec

Title: US-10-030-194A-5

Perfect score: 28

Sequence: 1 GYXVEE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : PIR\_79.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	96.4	74	JU0240	nuclear matrix pro
2	27	96.4	122	B87323	chemotaxis protein
3	27	96.4	233	G75075	hypothetical prote
4	27	96.4	280	WMBP8H	gene 8.5 protein -
5	27	96.4	280	WMBP8H	gene 8.5 protein -
6	27	96.4	313	A70441	glycerol-3-phospha
7	27	96.4	344	YFBSA	phenylalanine-tRNA
8	27	96.4	344	B83605	probable acetylpo
9	27	96.4	345	A40990	GTP-binding regula
10	27	96.4	363	A72702	hypothetical prote
11	27	96.4	390	T22810	hypothetical prote
12	27	96.4	395	T20912	CMP-N-acetylneuram
13	27	96.4	398	D90608	nadh-dependent fla
14	27	96.4	457	A57506	palindrome-binding
15	27	96.4	466	C86739	hypothetical prote
16	27	96.4	553	F90032	urocanate hydratase
17	27	96.4	646	AB2168	two-component hybr
18	27	96.4	858	A44919	GCR3 protein - yea
19	27	96.4	879	AC2542	cation-transportin
20	27	96.4	1374	A84888	hypothetical prote
21	26	92.9	80	C75026	hypothetical prote
22	26	92.9	95	E64631	hypothetical prote
23	26	92.9	95	C71883	hypothetical prote
24	26	92.9	116	AE1087	hypothetical prote
25	26	92.9	116	AD1451	hypothetical prote
26	26	92.9	124	C82805	DNA-binding protei
27	26	92.9	137	T45079	hypothetical prote
28	26	92.9	151	C71113	Ni,Fe-Hydrogenase
29	26	92.9	161	AG1132	hypothetical prote

30	26	92.9	177	2	T01785	protoporphyrin IX
31	26	92.9	181	2	T13309	hypothetical prote
32	26	92.9	204	2	H72256	conserved hypothet
33	26	92.9	207	2	S39103	ubiquinol-cytochro
34	26	92.9	242	2	AH1180	amino acid ABC tra
35	26	92.9	242	2	AI1537	amino acid ABC tra
36	26	92.9	245	1	NDECR5	type II site-speci
37	26	92.9	249	2	T06441	storage protein ho
38	26	92.9	252	2	C83837	hypothetical prote
39	26	92.9	254	1	UESY27	vegetative storage
40	26	92.9	257	2	H84713	hypothetical prote
41	26	92.9	275	2	S75190	hypothetical prote
42	26	92.9	285	2	H69802	conserved hypothet
43	26	92.9	295	2	AD2887	dioxygenase [impor
44	26	92.9	295	2	A97663	hypothetical prote
45	26	92.9	298	2	AH0947	probable sugar kin
46	26	92.9	313	2	H71438	hypothetical prote
47	26	92.9	315	2	S76043	hypothetical prote
48	26	92.9	319	2	S03833	hypothetical prote
49	26	92.9	326	2	S74349	hypothetical prote
50	26	92.9	337	2	AE1920	GTP-binding protei
51	26	92.9	354	2	G64475	GRP-binding protei
52	26	92.9	354	2	T52401	branched-chain ami
53	26	92.9	356	2	D96537	hypothetical prote
54	26	92.9	357	2	B47411	ADPribosylarginine
55	26	92.9	358	2	D89823	hypothetical prote
56	26	92.9	359	2	AF3184	beta-lactamase [im
57	26	92.9	361	2	T12613	adenosylmethionine
58	26	92.9	369	2	B84542	hypothetical prote
59	26	92.9	384	2	T00625	branched-chain ami
60	26	92.9	413	2	T43170	probable triacylg
61	26	92.9	413	2	G84462	probable SCARECROW
62	26	92.9	429	2	S46801	hypothetical prote
63	26	92.9	431	2	AE2090	site-specific DNA-
64	26	92.9	443	2	T39540	triglyceride lipas
65	26	92.9	483	2	E71133	hypothetical prote
66	26	92.9	511	2	G96688	hypothetical prote
67	26	92.9	523	2	T51475	RGA-like protein -
68	26	92.9	533	2	H86282	protein F10B6.34 [
69	26	92.9	575	2	S57648	nicotinic acetylch
70	26	92.9	576	2	T12896	probable single-st
71	26	92.9	587	2	D84426	hypothetical prote
72	26	92.9	606	2	T27072	hypothetical prote
73	26	92.9	611	2	F70325	conserved hypothet
74	26	92.9	629	2	E69284	signal-transducing
75	26	92.9	739	2	I40715	malate synthase [E
76	26	92.9	773	2	J02187	p87 protein - Card
77	26	92.9	836	2	D97505	copper-transportin
78	26	92.9	836	2	AH2723	heavy-metal transp
79	26	92.9	863	2	B69301	signal-transducing
80	26	92.9	873	1	A49729	VLDL receptor prec
81	26	92.9	898	2	H84701	probable villin li
82	26	92.9	900	2	G96617	probable disease x
83	26	92.9	910	2	T50671	villin 1 [imported
84	26	92.9	1010	2	T16616	hypothetical prote
85	26	92.9	1021	2	AC2202	hypothetical prote
86	26	92.9	1051	2	JC4091	glycoprotein A - P
87	26	92.9	1062	2	H83966	carbamoyl-phosphat
88	26	92.9	1087	1	S41797	cellulose 1,4-beta
89	26	92.9	1231	1	A48490	endo-1,4-beta-xyla
90	26	92.9	1234	2	S72640	endo-1,4-beta-xyla
91	26	92.9	1329	2	C69048	cobalamin biosynth
92	26	92.9	1379	2	S37310	protoporphyrin IX
93	26	92.9	1380	2	S64721	protoporphyrin IX
94	26	92.9	1381	2	S71288	protoporphyrin IX
95	26	92.9	1382	2	T01789	magnesium chelatase
96	26	92.9	1383	2	T07126	hypothetical prote
97	26	92.9	1653	2	T14758	multidrug resistanc
98	26	92.9	1743	2	T18279	multidrug resistanc
99	26	92.9	1905	2	T18267	hypothetical prote
100	25	89.3	144	2	G72267	hypothetical prote
101	25	89.3	149	2	S39910	S-receptor kinase
102	25	89.3	150	2	S31411	S-receptor kinase

103	25	89.3	180	2	F71809	hypothetical prote	176	24	85.7	244	2	AB2505	histidine kinase-1
104	25	89.3	180	2	B64711	purine nucleoside	177	24	85.7	251	2	G75063	hypothetical prote
105	25	89.3	202	2	D86776	amidotransferase (	178	24	85.7	255	2	JG0179	superoxide dismuta
106	25	89.3	202	2	I45734	imidazoleglycerol-	179	24	85.7	256	2	T41185	gene 38 protein -
107	25	89.3	240	2	H98031	hypothetical prote	180	24	85.7	262	2	T51732	superoxide dismuta
108	25	89.3	244	2	B95166	ABC transporter, A	181	24	85.7	296	1	S55795	arginase [EC 3.5.3
109	25	89.3	246	2	H86338	protein F2D10.6 [i	182	24	85.7	305	2	A69122	glutamine PRPP ami
110	25	89.3	249	2	H75174	ATP-binding protei	183	24	85.7	310	2	H71278	glutamine PRPP ami
111	25	89.3	249	2	G71039	probable PHNP prot	184	24	85.7	315	2	G84936	conserved hypothet
112	25	89.3	255	2	E70166	exodeoxyribonuclea	185	24	85.7	334	2	AE0154	kinase [imported]
113	25	89.3	266	2	T34172	hypothetical prote	186	24	85.7	350	1	RGBOT1	probable nucleoid-
114	25	89.3	267	2	B70350	hypothetical prote	187	24	85.7	350	1	RGHUT1	GTP-binding regula
115	25	89.3	286	1	S12797	ribosomal protein	188	24	85.7	350	1	RGHUT1	GTP-binding regula
116	25	89.3	300	2	S34116	transcription fact	189	24	85.7	351	2	E97225	carbamoylphosphate
117	25	89.3	300	2	T46883	transcription init	190	24	85.7	361	2	H87700	citrate synthase [
118	25	89.3	300	2	E75110	transcription init	191	24	85.7	366	2	AH2237	citrate synthase [
119	25	89.3	300	2	E71023	probable transcrip	192	24	85.7	370	1	QOHS4C	septom site-determ
120	25	89.3	311	2	B64998	hypothetical prote	193	24	85.7	370	1	QOHS4C	GTP-binding protei
121	25	89.3	311	2	E85867	hypothetical prote	194	24	85.7	370	1	QOHS4C	GTP-binding protei
122	25	89.3	311	2	D91023	hypothetical prote	195	24	85.7	370	1	QOHS4C	GTP-binding protei
123	25	89.3	315	2	S76596	hypothetical prote	196	24	85.7	370	1	QOHS4C	GTP-binding protei
124	25	89.3	324	2	E83082	probable ATP-bind	197	24	85.7	378	2	AB2089	2-hydroxyglutaryl-
125	25	89.3	331	2	AE1118	dihydroxyacetone k	198	24	85.7	385	1	YKYT	sulfolipid sulfoqu
126	25	89.3	331	2	AG1478	dihydroxyacetone k	199	24	85.7	385	1	YKYT	citrate (si)-synth
127	25	89.3	333	2	F90225	alcohol dehydrogen	200	24	85.7	387	2	H69080	aspartate transami
128	25	89.3	336	2	G90421	alcohol dehydrogen	201	24	85.7	401	2	AG3552	aspartate transami
129	25	89.3	339	2	H97107	uncharacterized pr	202	24	85.7	402	2	AG3552	branched-chain ami
130	25	89.3	360	2	T43254	GTP-binding protei	203	24	85.7	402	2	AG3552	hypothetical prote
131	25	89.3	376	1	E69957	gamma-D-glutamyl-L	204	24	85.7	428	2	B64301	amide-urea binding
132	25	89.3	379	2	AC0759	cblD protein limpo	205	24	85.7	428	2	B64301	hypothetical prote
133	25	89.3	383	2	AE1646	hypothetical prote	206	24	85.7	428	2	B64301	hypothetical prote
134	25	89.3	399	2	B64488	hypothetical prote	207	24	85.7	428	2	B64301	2,3-bisphosphoglyc
135	25	89.3	419	2	T37577	hypothetical prote	208	24	85.7	428	2	B64301	glutamate synthase
136	25	89.3	432	2	F81320	transcription term	209	24	85.7	428	2	B64301	membrane protein Y
137	25	89.3	434	2	G64444	amidase - Mechanoc	210	24	85.7	428	2	B64301	probable citrate s
138	25	89.3	469	2	T34173	hypothetical prote	211	24	85.7	522	2	T37536	citrate synthase-1
139	25	89.3	480	2	T49157	citrate synthase-1	212	24	85.7	522	2	T37536	cell polarity prot
140	25	89.3	565	1	HMIVGM	hemagglutinin prec	213	24	85.7	553	2	F90821	probable large ter
141	25	89.3	565	1	HMIVT3	hemagglutinin prec	214	24	85.7	553	2	F90821	probable terminase
142	25	89.3	566	1	HMIVT2	hemagglutinin prec	215	24	85.7	553	2	F90821	probable terminase
143	25	89.3	566	1	HMIVT1	hemagglutinin prec	216	24	85.7	553	2	F90821	probable terminase
144	25	89.3	681	2	E64598	hypothetical prote	217	24	85.7	553	2	F90821	unknown protein en
145	25	89.3	683	2	H71914	hypothetical prote	218	24	85.7	553	2	F90821	carbamoyl-phosphat
146	25	89.3	684	2	E64496	ATP-dependent RNA	219	24	85.7	553	2	F90821	hypothetical prote
147	25	89.3	865	1	ISECTP	DNA topoisomerase	220	24	85.7	553	2	F90821	conserved hypothet
148	25	89.3	865	2	AP0654	DNA topoisomerase	221	24	85.7	553	2	F90821	phenylalanine-tRNA
149	25	89.3	865	2	F90859	DNA topoisomerase	222	24	85.7	553	2	F90821	phenylalanine-tRNA
150	25	89.3	865	2	A85760	DNA topoisomerase	223	24	85.7	553	2	F90821	centromere protein
151	25	89.3	868	2	D83269	DNA topoisomerase	224	24	85.7	553	2	F90821	lipoxigenase (EC 1
152	25	89.3	868	2	E64119	hypothetical prote	225	24	85.7	553	2	F90821	lipoxigenase (EC 1
153	25	89.3	868	2	E90308	NAD(P) transhydrog	226	24	85.7	553	2	F90821	lipoxigenase (EC 1
154	25	89.3	1086	2	S54876	unknown protein Tl	227	24	85.7	553	2	F90821	lipoxigenase (EC 1
155	25	89.3	1088	2	H96747	hypothetical prote	228	24	85.7	553	2	F90821	lipoxigenase (EC 1
156	25	89.3	1119	2	T32074	hypothetical prote	229	24	85.7	553	2	F90821	lipoxigenase (EC 1
157	25	89.3	1260	2	A72603	probable nitrate r	230	24	85.7	553	2	F90821	lipoxigenase (EC 1
158	25	89.3	1328	2	AE2351	protoporphyrin IX	231	24	85.7	553	2	F90821	lipoxigenase (EC 1
159	25	89.3	1331	2	S75000	protoporphyrin IX	232	24	85.7	553	2	F90821	lipoxigenase (EC 1
160	25	89.3	1768	2	T27023	hypothetical prote	233	24	85.7	553	2	F90821	lipoxigenase (EC 1
161	25	89.3	1983	2	AC1922	two-component hybr	234	24	85.7	553	2	F90821	lipoxigenase (EC 1
162	25	89.3	2413	2	S34670	splicing factor PR	235	24	85.7	553	2	F90821	lipoxigenase (EC 1
163	24	85.7	72	2	D83830	hypothetical prote	236	24	85.7	553	2	F90821	lipoxigenase (EC 1
164	24	85.7	74	2	S06777	hypothetical prote	237	24	85.7	553	2	F90821	lipoxigenase (EC 1
165	24	85.7	76	2	AI3049	transcription regu	238	24	85.7	553	2	F90821	lipoxigenase (EC 1
166	24	85.7	119	2	B98236	exs1 protein prote	239	24	85.7	553	2	F90821	lipoxigenase (EC 1
167	24	85.7	133	2	B69489	ribosomal protein	240	24	85.7	553	2	F90821	lipoxigenase (EC 1
168	24	85.7	136	2	E69179	ribosomal protein	241	24	85.7	553	2	F90821	lipoxigenase (EC 1
169	24	85.7	152	2	E64322	probable secreted	242	24	85.7	553	2	F90821	lipoxigenase (EC 1
170	24	85.7	168	2	T35779	hypothetical prote	243	24	85.7	553	2	F90821	lipoxigenase (EC 1
171	24	85.7	179	2	D72459	probable secreted	244	24	85.7	553	2	F90821	lipoxigenase (EC 1
172	24	85.7	180	2	T34851	hypothetical prote	245	24	85.7	553	2	F90821	lipoxigenase (EC 1
173	24	85.7	189	2	D83008	outer membrane lip	246	24	85.7	553	2	F90821	lipoxigenase (EC 1
174	24	85.7	211	2	G70981	hypothetical prote	247	24	85.7	553	2	F90821	lipoxigenase (EC 1
175	24	85.7	217	2	H86781	metalloreceptor l	248	24	85.7	553	2	F90821	lipoxigenase (EC 1



249	23	82.1	83	2	B95974	probable transcrip	322	23	82.1	256	1	S30826	H+-exporting ATPase
250	23	82.1	84	2	H69092	conserved hypotnet	323	23	82.1	261	2	A51471	hypothetical prote
251	23	82.1	85	2	A12111	hypothetical prote	324	23	82.1	261	2	H98213	transcription repr
252	23	82.1	86	2	G86727	hypothetical prote	325	23	82.1	272	2	B70354	hypothetical prote
253	23	82.1	96	2	H66631	hypothetical prote	326	23	82.1	275	2	B7474	enoyl-CoA hydratase
254	23	82.1	96	2	H86641	hypothetical prote	327	23	82.1	283	1	B69547	iron-sulfur cluste
255	23	82.1	99	2	A82651	hypothetical prote	328	23	82.1	285	2	AG3306	dipeptide transpor
256	23	82.1	103	1	R8882B	ribosomal protein	329	23	82.1	286	2	B90156	hypothetical prote
257	23	82.1	105	2	T45377	ribosomal protein	330	23	82.1	287	2	T46874	ribokinase homolog
258	23	82.1	105	2	F70643	probable ribosomal	331	23	82.1	288	2	T21732	hypothetical prote
259	23	82.1	111	2	T17132	hypothetical prote	332	23	82.1	288	2	T15735	hypothetical prote
260	23	82.1	113	2	C70691	hypothetical prote	333	23	82.1	289	2	T40596	probable DNA repai
261	23	82.1	121	2	D83624	probable two-compo	334	23	82.1	291	2	A72341	hypothetical prote
262	23	82.1	127	2	H5146	hypothetical prote	335	23	82.1	292	2	G81944	probable signal pe
263	23	82.1	138	2	T46133	hypothetical prote	336	23	82.1	294	2	T27791	hypothetical prote
264	23	82.1	143	2	D71015	hypothetical prote	337	23	82.1	296	2	AB0557	cytochrome o ubiqu
265	23	82.1	143	2	E71129	hypothetical prote	338	23	82.1	303	2	H69551	acyl carrier prote
266	23	82.1	144	2	B72280	hypothetical prote	339	23	82.1	304	2	T23588	hypothetical prote
267	23	82.1	147	2	F69124	hypothetical prote	340	23	82.1	307	2	C69664	conserved hypotnet
268	23	82.1	150	2	H87682	Maoc family protei	341	23	82.1	307	2	T04141	histone deacetylase
269	23	82.1	151	2	A96948	peptidoglycan-bind	342	23	82.1	308	2	D90261	conserved hypotnet
270	23	82.1	157	2	T21699	hypothetical prote	343	23	82.1	309	2	E83999	mutants block spor
271	23	82.1	162	2	D69899	conserved hypotnet	344	23	82.1	311	2	S77803	hypothetical prote
272	23	82.1	163	2	C71547	hypothetical prote	345	23	82.1	313	2	F65059	hypothetical prote
273	23	82.1	164	2	G82567	transcription regu	346	23	82.1	316	2	AC1273	ornithine carbamoy
274	23	82.1	164	2	A81325	probable signal-tr	347	23	82.1	316	2	AD1636	ornithine carbamoy
275	23	82.1	166	2	A98310	probable transcrip	348	23	82.1	317	2	AH2729	agmatinase [impor
276	23	82.1	168	2	AB2973	sigma factor [impo	349	23	82.1	317	2	A97511	hypothetical prote
277	23	82.1	168	2	F95941	probable small hea	350	23	82.1	324	2	B75163	anthranilate synth
278	23	82.1	170	2	S19687	hypothetical prote	351	23	82.1	325	2	T31977	hypothetical prote
279	23	82.1	176	2	T43335	vacuolar sorting p	352	23	82.1	327	2	S62817	DNA-directed RNA p
280	23	82.1	179	2	T41274	probable translati	353	23	82.1	327	2	G72411	hypothetical prote
281	23	82.1	180	2	D70161	ribosomal protein	354	23	82.1	328	2	A93848	ferric ion ABC tra
282	23	82.1	185	2	T12772	conserved hypotnet	355	23	82.1	331	2	F64219	DNA-directed RNA p
283	23	82.1	192	2	T41460	hypothetical 51.9	356	23	82.1	331	2	E82389	probable outer mem
284	23	82.1	194	2	S45561	RNA polymerase sig	357	23	82.1	332	2	I40792	hypothetical prote
285	23	82.1	195	2	S28845	myosin regulatory	358	23	82.1	333	2	A82126	conserved hypotnet
286	23	82.1	197	2	A90365	hypothetical prote	359	23	82.1	335	2	AC0786	nucleoid-associated
287	23	82.1	197	2	S42132	light-harvesting c	360	23	82.1	335	2	A64988	hypothetical 37.8
288	23	82.1	197	2	S42133	light-harvesting c	361	23	82.1	335	2	H85857	nucleoid-associate
289	23	82.1	197	2	S42129	light-harvesting c	362	23	82.1	335	2	F91013	nucleoid-associate
290	23	82.1	197	2	S42130	light-harvesting c	363	23	82.1	336	2	S69524	hypothetical prote
291	23	82.1	202	2	G87352	hypothetical prote	364	23	82.1	336	2	T04024	hypothetical prote
292	23	82.1	206	2	AC2443	orotate phosphorib	365	23	82.1	337	1	NPBS	transcription regu
293	23	82.1	209	2	AH2716	NTP pyrophosphohy	366	23	82.1	337	2	G86694	C4-dicarboxylate t
294	23	82.1	209	2	E96517	hypothetical prote	367	23	82.1	339	2	B81162	ccsA protein - Eug
295	23	82.1	211	2	I77569	tet repressor - Es	368	23	82.1	341	2	A11429	LPS biosynthesis R
296	23	82.1	214	1	RGBSCA	regulatory protein	369	23	82.1	341	2	AG1803	Phe-tRNA synthetas
297	23	82.1	215	2	AG3072	transcription regu	370	23	82.1	342	2	B81929	probable iron-upta
298	23	82.1	215	2	E75058	hypothetical prote	371	23	82.1	348	2	F83993	unknown protein T2
299	23	82.1	216	2	S59942	response regulator	372	23	82.1	348	2	S34494	site-specific DNA-
300	23	82.1	217	2	F97085	response regulator	373	23	82.1	350	2	D69143	probable N-acetyl-
301	23	82.1	218	2	T43438	hypothetical prote	374	23	82.1	352	2	F89884	iron utilization p
302	23	82.1	221	2	C81155	NAD(P)H nitroreduc	375	23	82.1	352	2	B81929	nucleotide-binding
303	23	82.1	221	2	E81949	probable NAD(P)H-f	376	23	82.1	352	2	F86836	hypothetical prote
304	23	82.1	222	2	S47066	hypothetical prote	377	23	82.1	355	1	C64526	cysteine proteinase
305	23	82.1	225	2	T26209	hypothetical prote	378	23	82.1	355	2	F72625	similar to YeeB (B
306	23	82.1	225	2	AG1397	two-component resp	379	23	82.1	356	2	T10888	hypothetical prote
307	23	82.1	225	2	AG1773	two-component resp	380	23	82.1	356	2	B64048	hypothetical prote
308	23	82.1	226	2	G83293	probable two-compo	381	23	82.1	359	2	G86290	hypothetical prote
309	23	82.1	227	2	A72746	probable ribonucle	382	23	82.1	360	2	S59538	hypothetical prote
310	23	82.1	227	2	AB2008	hypothetical prote	383	23	82.1	361	2	H96943	ABC transporter, A
311	23	82.1	231	2	T02765	glutathione transf	384	23	82.1	362	2	T20222	carbamoyl-phosphat
312	23	82.1	233	2	H69130	conserved hypotnet	385	23	82.1	362	2	F90441	carbamoyl-phosphat
313	23	82.1	234	2	G95008	phosphorylase, Pnp	386	23	82.1	363	2	AD1304	cofactor biosynth
314	23	82.1	234	2	C69519	conserved hypotnet	387	23	82.1	363	2	AD1676	hypothetical prote
315	23	82.1	239	2	E84044	two-component resp	388	23	82.1	364	1	H69207	hypothetical prote
316	23	82.1	240	1	RGBSAP	phosphate response	389	23	82.1	364	2	T22807	hypothetical prote
317	23	82.1	245	2	F96913	glutamine ABC tran	390	23	82.1	367	2	C96537	hypothetical prote
318	23	82.1	251	2	A83917	transcription regu	391	23	82.1	367	2	T32141	hypothetical prote
319	23	82.1	251	2	E96637	hypothetical prote	392	23	82.1	368	2	T44887	probable branched-
320	23	82.1	254	2	D97880	conserved hypotnet	393	23	82.1	368	2	C70786	probable ilvE prot
321	23	82.1	254	2	C97498	hypothetical 21.7K	394	23	82.1	368	2	A11934	hypothetical prote

335	23	82.1	371	2	D83983	hippurate hydrolas	468	23	82.1	516	2	T52611	glucose-6-phosphat
336	23	82.1	372	2	T00243	sopA protein - Bsc	469	23	82.1	517	2	G83311	probable transcrip
337	23	82.1	373	2	E86344	hypothetical prote	470	23	82.1	519	2	A82634	2-isopropylmalate
338	23	82.1	376	2	B84360	citrate synthase I	471	23	82.1	530	2	E35116	anthranilate synth
339	23	82.1	377	2	T47471	cysteine proteinas	472	23	82.1	530	2	B83519	sigma-54 depend
340	23	82.1	378	2	E70786	hypothetical prote	473	23	82.1	530	2	E82491	glucuronosyltransf
341	23	82.1	379	2	E71296	probable spermidin	474	23	82.1	530	2	S68200	hypothetical prote
342	23	82.1	382	2	E82249	mrp protein VC1037	475	23	82.1	530	2	S38903	signal-transducing
343	23	82.1	384	2	S14450	probable transposa	476	23	82.1	530	2	T02995	unspecific monooxy
344	23	82.1	386	2	A70740	probable fadE14 pr	477	23	82.1	531	2	B69346	probable fadE8 pro
345	23	82.1	388	1	BVECAF	sopA protein - Bsc	478	23	82.1	542	2	A70826	signal-transducing
346	23	82.1	388	2	D91240	probable amino aci	479	23	82.1	547	2	S70538	methylmalonyl-CoA
347	23	82.1	388	2	A86088	probable hippurica	480	23	82.1	548	1	G69526	hypothetical prote
348	23	82.1	389	2	AF1251	Acetyl-CoA acetyl	481	23	82.1	548	2	AG2931	hypothetical prote
349	23	82.1	394	2	E72553	probable citrate s	482	23	82.1	555	2	S71365	ovule development
350	23	82.1	394	2	A82243	acetate kinase Vcl	483	23	82.1	556	2	C72204	alpha-amylase - Th
351	23	82.1	398	2	A82243	ATP-gated ion chan	484	23	82.1	560	2	T52438	PRM1 homolog [impo
352	23	82.1	399	2	D89797	hypothetical prote	485	23	82.1	562	2	T49386	hypothetical prote
353	23	82.1	399	2	A83529	chemotaxis motD pr	486	23	82.1	570	2	B85056	probable transposo
354	23	82.1	399	2	A83529	protein W10D9.4 [i	487	23	82.1	571	2	C70762	hypothetical ABC t
355	23	82.1	403	2	E88021	hypothetical prote	488	23	82.1	575	2	G98350	sensor histidine k
356	23	82.1	403	2	D70529	hypothetical 44.6	489	23	82.1	577	2	B75621	uncharacterized pr
357	23	82.1	408	1	G65132	beta-2-adrenergic	490	23	82.1	581	2	C96991	2,4-dichlorophenol
358	23	82.1	413	1	QRHUB2	sensory transducti	491	23	82.1	588	2	E35255	FOG2 protein- yea
359	23	82.1	413	2	A69205	hypothetical prote	492	23	82.1	602	2	S72513	aspartate-tRNA lig
360	23	82.1	417	2	G69250	hypothetical prote	493	23	82.1	605	2	F71724	probable tpr prote
361	23	82.1	419	2	T19260	hypothetical prote	494	23	82.1	608	2	H71379	hypothetical prote
362	23	82.1	421	2	E64329	hypothetical prote	495	23	82.1	613	2	T52465	hypothetical prote
363	23	82.1	423	2	A69367	translation elonga	496	23	82.1	613	2	T27077	hypothetical prote
364	23	82.1	423	2	H70808	probable lppF prot	497	23	82.1	639	2	T23558	hypothetical prote
365	23	82.1	430	2	D72411	folypolyglutamate	498	23	82.1	662	2	E95105	ABC transporter, p
366	23	82.1	438	2	C82203	methyl-accepting c	499	23	82.1	674	1	S69731	ABC transporter, A
367	23	82.1	438	2	T37229	hypothetical prote	500	23	82.1	677	2	F75321	serine/threonine-s
368	23	82.1	440	2	T24073	hypothetical prote	501	23	82.1	677	2	T45921	probable NAD synth
369	23	82.1	441	2	E69454	ribulose-bisphosph	502	23	82.1	680	2	H87091	hypothetical prote
370	23	82.1	442	2	H72266	astB/chnr-related	503	23	82.1	681	2	T01469	prolyl oligopeptid
371	23	82.1	443	2	H75447	thiophene and fura	504	23	82.1	683	2	E87495	hypothetical prote
372	23	82.1	443	2	F89426	protein M162.5 [im	505	23	82.1	694	2	T00148	glycerol kinase (E
373	23	82.1	443	2	T04025	hypothetical prote	506	23	82.1	709	2	S33907	allanyl dipeptidyl
374	23	82.1	446	2	T45525	conserved hypothet	507	23	82.1	709	2	B82580	pilQ protein - Pse
375	23	82.1	446	2	F85277	WSC4 homolog limpo	508	23	82.1	714	2	S37345	type 4 fibrial bi
376	23	82.1	448	2	H95090	probable glutamate	509	23	82.1	714	2	A83016	TN916 ORF15 homolo
377	23	82.1	448	2	D97958	glutathione-disulf	510	23	82.1	719	2	AI1212	hypothetical prote
378	23	82.1	449	2	AB1188	glutathione-disulf	511	23	82.1	723	2	T32136	hypothetical prote
379	23	82.1	449	2	AB1546	glutathione-disulf	512	23	82.1	738	2	D70680	isoamylase (EC 3.2
380	23	82.1	449	2	S41386	glutathione-disulf	513	23	82.1	776	2	S13470	isoamylase (EC 3.2
381	23	82.1	450	2	T39443	probable triglycer	514	23	82.1	776	2	A37035	signal-transducing
382	23	82.1	460	2	T21678	hypothetical prote	515	23	82.1	781	2	C69452	glycerophosphoryl
383	23	82.1	460	2	G71082	probable mannose-1	516	23	82.1	785	2	D75630	cation-transportin
384	23	82.1	464	2	T00956	translation initia	517	23	82.1	813	2	AH3258	endopeptidase La (
385	23	82.1	465	1	G75104	methyl-accepting c	518	23	82.1	819	1	C71527	endopeptidase La (
386	23	82.1	465	2	A82407	triglyceride lipas	519	23	82.1	819	1	E86494	Lon ATP-dependent
387	23	82.1	467	2	T41053	transcription regu	520	23	82.1	819	2	E81681	proteinase, Lon fa
388	23	82.1	469	2	B83722	PRM1 homolog [impo	521	23	82.1	819	2	T45633	starch phosphoryla
389	23	82.1	470	2	AD2440	hypothetical prote	522	23	82.1	844	2	S77547	ethylene response
390	23	82.1	470	2	AD2435	probable glutamyl-	523	23	82.1	853	2	A95269	probable sensory t
391	23	82.1	471	1	T40579	isochoismate synt	524	23	82.1	864	2	F94176	bacterio-opsin act
392	23	82.1	474	2	S75568	conserved hypothet	525	23	82.1	867	1	RRVECV	RNA-directed RNA p
393	23	82.1	474	2	E72265	polysaccharide bio	526	23	82.1	873	1	I48952	VLDL receptor prec
394	23	82.1	474	2	AB2990	polysaccharide bio	527	23	82.1	873	1	QRRBYD	VLDL receptor prec
395	23	82.1	479	1	TVCH22	transcription fact	528	23	82.1	873	2	AB1944	hypothetical prote
396	23	82.1	481	2	S76820	hypothetical prote	529	23	82.1	885	2	T47966	hypothetical prote
397	23	82.1	485	2	T01968	isoleucine-tRNA li	530	23	82.1	902	2	D71079	probable disease r
398	23	82.1	486	2	S30959	gene 14 protein -	531	23	82.1	902	2	F96617	probable two-compo
399	23	82.1	486	2	T23776	hypothetical prote	532	23	82.1	907	2	H95973	probable two-compo
400	23	82.1	487	2	AI1536	hypothetical prote	533	23	82.1	915	1	A48225	subtilisin-like pr
401	23	82.1	489	2	B31944	cytochrome P450 2G	534	23	82.1	915	2	B48225	probable propotei
402	23	82.1	494	2	A35551	cytochrome P450 2G	535	23	82.1	939	2	T32521	hypothetical prote
403	23	82.1	494	2	B69731	PASX prophage ORF	536	23	82.1	965	2	A42452	two-component hybr
404	23	82.1	495	2	D97023	threonine synthase	537	23	82.1	974	2	AC2076	two-component hybr
405	23	82.1	496	2	D97023	hypothetical prote	538	23	82.1	1017	2	T24349	myosin IA - Caenor
406	23	82.1	507	2	T00662	glucose-6-phosphat	539	23	82.1	1049	2	B90124	second-largest sub
407	23	82.1	515	2	T52610		540	23	82.1				

541	23	82.1	1085	2	JC2227	probable helicase	614	22	78.6	123	2	C96604	hypothetical prote
542	23	82.1	1086	2	A88855	protein M18.5 [imp	615	22	78.6	124	2	F69135	conserved hypothet
543	23	82.1	1089	2	E81446	carbamoyl-phosphat	616	22	78.6	127	2	G71186	hypothetical prote
544	23	82.1	1094	2	T05472	hypothetical prote	617	22	78.6	129	2	G90761	hypothetical prote
545	23	82.1	1134	2	T23798	hypothetical prote	618	22	78.6	129	2	B85625	hypothetical prote
546	23	82.1	1156	2	A47397	adducin homolog -	619	22	78.6	131	2	D84401	30S ribosomal prot
547	23	82.1	1197	1	G65010	sensor protein evg	620	22	78.6	131	2	B69527	conserved hypothet
548	23	82.1	1197	2	A91035	probable sensor fo	621	22	78.6	133	2	E75215	glyoxalase I relat
549	23	82.1	1229	2	B95879	probable sensor fo	622	22	78.6	136	2	H83194	hypothetical prote
550	23	82.1	1299	2	B95987	probable two-compo	623	22	78.6	136	2	D75265	conserved hypothet
551	23	82.1	1230	2	S53974	hypothetical prote	624	22	78.6	136	2	A71452	glyoxalase I relat
552	23	82.1	1298	2	T47523	DNA-binding protei	625	22	78.6	137	2	JH0433	transformation com
553	23	82.1	1320	2	E59092	hypothetical prote	626	22	78.6	137	2	G59473	hypothetical prote
554	23	82.1	1343	1	H64073	DNA-directed RNA p	627	22	78.6	138	2	T49060	hypothetical prote
555	23	82.1	1357	2	H83112	DNA-directed RNA p	628	22	78.6	142	2	H72520	hypothetical prote
556	23	82.1	1437	2	C75198	activator 1, repli	629	22	78.6	143	2	H82363	conserved hypothet
557	23	82.1	1489	2	T31108	cyst germination s	630	22	78.6	148	2	AF1081	50S ribosomal prot
558	23	82.1	1548	2	S34583	serine proteinase	631	22	78.6	148	2	AG1438	50S ribosomal prot
559	23	82.1	1583	2	T14176	probable phosphati	632	22	78.6	148	2	AD0483	universal stress p
560	23	82.1	1585	2	T31611	hypothetical prote	633	22	78.6	148	2	G75086	hypothetical prote
561	23	82.1	1607	2	G87259	conserved hypothet	634	22	78.6	150	2	H82150	hypothetical prote
562	23	82.1	1613	2	A43081	vitellogenin vit-2	635	22	78.6	151	2	B71166	hypothetical prote
563	23	82.1	1613	2	F89528	protein vit-2 [imp	636	22	78.6	152	2	B64485	hypothetical prote
564	23	82.1	1627	2	A82109	two-component hybr	637	22	78.6	153	2	B96959	anaerobic ribonuc
565	23	82.1	1645	2	T31339	carbamoyl-phosphat	638	22	78.6	154	2	S19721	pyruvate dehydroge
566	23	82.1	1666	2	T43169	hypothetical prote	639	22	78.6	154	2	T39892	probable prefolidin
567	23	82.1	1926	2	JC4842	DNA-binding nuclea	640	22	78.6	155	2	T13308	hypothetical prote
568	23	82.1	2332	1	GNNY2F	genome polypotein	641	22	78.6	157	2	PH0201	hypothetical prote
569	23	82.1	2333	1	GNNY2F	genome polypotein	642	22	78.6	158	2	D83384	probable transcrip
570	23	82.1	2581	2	A25545	hypothetical prote	643	22	78.6	158	2	A55025	hypothetical prote
571	23	82.1	2688	2	I49477	alpha-A-crystallin	644	22	78.6	158	2	D85892	hypothetical prote
572	23	82.1	2717	2	A34203	DNA-binding protei	645	22	78.6	162	2	H91047	hypothetical prote
573	23	82.1	3386	1	GNWDF9	genome polypotein	646	22	78.6	162	2	H69504	hypothetical prote
574	23	82.1	6359	2	T31679	bacitracin synthet	647	22	78.6	163	2	S49633	hypothetical prote
575	23	82.1	6831	2	A98852	protein unc-22 [im	648	22	78.6	165	2	G69362	hypothetical prote
576	23	82.1	6839	2	S72424	twitchin [similar	649	22	78.6	167	2	A75214	hypothetical prote
577	23	82.1	7160	2	T27935	hypothetical prote	650	22	78.6	169	2	A64356	hypothetical prote
578	23	82.1	76926	1	I38344	titin, Cardiac mus	651	22	78.6	171	2	E98288	hypothetical prote
579	22	78.6	24	2	A47209	histone H1-MDBP-2	652	22	78.6	171	2	AE2995	hypothetical prote
580	22	78.6	36	2	C95218	conserved domain p	653	22	78.6	173	2	T42853	late expression fa
581	22	78.6	43	2	S24180	histone H1.a. test	654	22	78.6	173	2	T41773	LEF-6 orf28 - Bomb
582	22	78.6	57	2	D64502	hypothetical prote	655	22	78.6	174	2	H70702	hypothetical prote
583	22	78.6	60	2	B44776	hypothetical prote	656	22	78.6	174	2	S47089	finger protein H2P
584	22	78.6	64	1	NTSRIC	neurotoxin I - bar	657	22	78.6	174	2	A70395	hypothetical prote
585	22	78.6	65	2	AC2154	hypothetical prote	658	22	78.6	178	2	G71198	hypothetical prote
586	22	78.6	73	2	A38927	estrogen receptor	659	22	78.6	180	2	G72616	hypothetical prote
587	22	78.6	78	1	R0ECD5	dihydrofolate redu	660	22	78.6	180	2	T06718	hypothetical prote
588	22	78.6	78	1	R0ECD6	dihydrofolate redu	661	22	78.6	181	2	S59505	ferric pseudobacti
589	22	78.6	78	1	R0ECD8	dihydrofolate redu	662	22	78.6	181	2	F97275	histidinol phospho
590	22	78.6	78	2	T08517	dihydrofolate redu	663	22	78.6	182	2	E87289	conserved hypothet
591	22	78.6	78	2	S32183	dihydrofolate redu	664	22	78.6	183	2	E70588	hypothetical prote
592	22	78.6	79	2	S23924	histone H1.b. hepa	665	22	78.6	184	2	A45503	41-2 protein antig
593	22	78.6	81	1	IHTF	high potential iro	666	22	78.6	186	2	F64323	ribosomal protein
594	22	78.6	82	2	S68349	H+-transporting tw	667	22	78.6	187	2	AG2528	hypothetical prote
595	22	78.6	82	2	T07211	H+-transporting tw	668	22	78.6	188	2	D87705	intracellular sept
596	22	78.6	84	2	F72497	hypothetical prote	669	22	78.6	193	2	T08300	hypothetical prote
597	22	78.6	86	2	C97083	uncharacterized pr	670	22	78.6	194	1	HSTR1	histone H1 - trout
598	22	78.6	89	2	S24178	hypothetical prote	671	22	78.6	196	2	S56961	probable membrane
599	22	78.6	93	2	T31048	hypothetical prote	672	22	78.6	198	1	C69986	conserved hypothet
600	22	78.6	97	2	C37202	hypothetical prote	673	22	78.6	198	2	AC1384	protein gp51 [Bact
601	22	78.6	98	2	H95213	hypothetical prote	674	22	78.6	203	2	F84191	hypothetical prote
602	22	78.6	98	2	A98078	hypothetical prote	675	22	78.6	206	1	HSTR1R	histone H1 - rainb
603	22	78.6	100	1	S40190	ribosomal protein	676	22	78.6	207	2	A56190	titin - rat (fragm
604	22	78.6	105	1	H8B011	histone H1.1 - bov	677	22	78.6	207	2	JH0550	histone Hit - huma
605	22	78.6	106	2	A72336	conserved hypothet	678	22	78.6	208	1	HSTR1T	histone Hit - rat
606	22	78.6	108	2	A64511	hypothetical prote	679	22	78.6	208	2	AF1504	amidotransferases
607	22	78.6	108	2	S64602	hypothetical prote	680	22	78.6	208	2	I70195	histone Hit - rhes
608	22	78.6	108	2	G64386	hypothetical prote	681	22	78.6	208	2	S43434	histone H1, testis
609	22	78.6	109	2	G71077	hypothetical prote	682	22	78.6	208	2	A41030	troponin I, cardia
610	22	78.6	119	2	T49848	hypothetical prote	683	22	78.6	209	2	A54984	ELF-1 protein prec
611	22	78.6	120	2	A69453	hypothetical prote	684	22	78.6	209	2	AB1444	gp51 (Bacteriophag
612	22	78.6	120	2	E83826	hypothetical prote	685	22	78.6	210	2	T06864	transcription regu
613	22	78.6	120	2	PC2201	hypothetical prote	686	22	78.6	210	2	AD0671	hypothetical prote

687	22	78.6	211	1	HSPG1T	histone H1t - pig	760	22	78.6	265	2	F70233	hypothetical prote
688	22	78.6	211	2	T01194	hypothetical prote	761	22	78.6	270	2	F70432	conserved hypothet
689	22	78.6	211	2	T32354	hypothetical prote	762	22	78.6	272	2	B95852	conserved hypothet
690	22	78.6	212	2	A28470	histone H1 - mouse	763	22	78.6	272	2	S33622	ADRF protein - soy
691	22	78.6	213	1	HSRU11	histone H1-2 [vali	764	22	78.6	275	2	E72575	probable 5'-methyl
692	22	78.6	213	1	HSRB13	histone H1.3 - rab	765	22	78.6	275	2	S77009	hypothetical prote
693	22	78.6	213	2	S43949	histone H1 - mouse	766	22	78.6	276	2	T08896	Sall3-2 protein, a
694	22	78.6	213	2	SJ03322	ephrlin-A2 - human	767	22	78.6	278	2	D71128	hypothetical prote
695	22	78.6	214	2	JQ2001	hypothetical 24K p	768	22	78.6	279	2	S03804	hypothetical prote
696	22	78.6	215	2	S26363	histone H1.1 - hum	769	22	78.6	284	2	AD0413	pantoate-beta-alan
697	22	78.6	215	2	AB2141	hypothetical prote	770	22	78.6	284	2	E70232	hypothetical prote
698	22	78.6	217	2	JH0159	histone H1d - rat	771	22	78.6	286	2	T20149	hypothetical prote
699	22	78.6	218	1	HSCH1	histone H1.02 - ch	772	22	78.6	286	2	B70722	hypothetical prote
700	22	78.6	218	2	S01262	histone H1 - musco	773	22	78.6	288	2	C75570	chromosome partiti
701	22	78.6	218	2	A23055	histone H1.01 - ch	774	22	78.6	289	2	T17957	hypothetical prote
702	22	78.6	219	1	HSU1B	histone H1-4 [vali	775	22	78.6	289	2	G86858	conserved hypothet
703	22	78.6	219	2	C28456	histone H1.11R - c	776	22	78.6	289	2	T41846	HE65 orf105 - Bomb
704	22	78.6	219	2	I49742	histone H1 - mouse	777	22	78.6	291	2	T08848	vegetative storage
705	22	78.6	220	2	A28456	histone H1.10 - ch	778	22	78.6	293	2	T27430	hypothetical prote
706	22	78.6	220	2	A72153	O2L protein - vari	779	22	78.6	295	2	S77308	hypothetical prote
707	22	78.6	220	2	H42504	M2L protein - vacc	780	22	78.6	295	2	AG3240	conserved hypothet
708	22	78.6	220	2	T28457	hypothetical prote	781	22	78.6	296	2	H87508	phosphoserine phos
709	22	78.6	220	2	I36838	O2L protein - vari	782	22	78.6	296	2	D89336	hypothetical prote
710	22	78.6	220	2	C94383	hypothetical prote	783	22	78.6	298	2	G81708	enoyl-(acyl)-carrie
711	22	78.6	221	1	B40335	histone H1-3 [vali	784	22	78.6	298	2	G71556	probable enoyl-lac
712	22	78.6	221	2	S49482	histone H1 - mouse	785	22	78.6	299	2	S75156	hypothetical prote
713	22	78.6	222	2	C72232	hypothetical prote	786	22	78.6	299	2	A69343	dihydroorotase deh
714	22	78.6	222	2	C67033	hypothetical prote	787	22	78.6	301	2	D83010	probable binding p
715	22	78.6	223	2	S49492	histone H1 - mouse	788	22	78.6	304	2	S33496	stereospecific hal
716	22	78.6	224	2	D28456	histone H1.03 - ch	789	22	78.6	304	2	T16828	trypsin (EC 3.4.21
717	22	78.6	225	2	C71721	ribonuclease III (	790	22	78.6	305	2	T52111	hypothetical prote
718	22	78.6	225	2	B28456	histone H1.11L - c	791	22	78.6	307	2	A72632	stearyl-CoA 9-des
719	22	78.6	226	1	S11660	histone H1-5 [vali	792	22	78.6	307	2	D70384	hypothetical prote
720	22	78.6	226	2	AC1757	hypothetical prote	793	22	78.6	310	1	S02137	beta lactamase pre
721	22	78.6	226	2	AE1650	hypothetical prote	794	22	78.6	310	2	G72752	carbamate kinase (
722	22	78.6	226	2	AF1589	hypothetical prote	795	22	78.6	311	2	A81179	probable sugar kin
723	22	78.6	228	2	F86930	probable two-compo	796	22	78.6	312	2	F64435	protein-export mem
724	22	78.6	229	2	AF3022	two component resp	797	22	78.6	312	2	C71136	mevalonate kinase
725	22	78.6	229	2	D98262	DNA-binding respon	798	22	78.6	312	2	E72516	hypothetical prote
726	22	78.6	229	2	H64367	dolichyl-phosphate	799	22	78.6	313	2	H25136	probable ribokinas
727	22	78.6	229	2	T40439	hypothetical prote	800	22	78.6	314	1	EQDVSG	cytochrome-c3 hydr
728	22	78.6	230	2	A70821	hypothetical prote	801	22	78.6	314	2	D70313	riboflavin kinase
729	22	78.6	230	2	G90073	conserved hypothet	802	22	78.6	315	2	D71012	dihydroxyacetone k
730	22	78.6	231	2	S45089	hypothetical prote	803	22	78.6	315	2	E70538	hypothetical prote
731	22	78.6	233	2	H82163	hypothetical prote	804	22	78.6	316	2	T50274	export protein xfo
732	22	78.6	233	2	AF2051	arginyl-tRNA-prote	805	22	78.6	316	2	E82809	hypothetical prote
733	22	78.6	236	2	T12810	two-component syst	806	22	78.6	319	2	A70638	probable ribose-ph
734	22	78.6	236	2	C95073	hypothetical prote	807	22	78.6	322	2	A72569	probable ATP-bindi
735	22	78.6	238	2	B97941	conserved hypothet	808	22	78.6	323	2	D72484	cathepsin - Orgyia
736	22	78.6	238	2	B84167	glucose-1-phosphat	809	22	78.6	324	2	T10394	cathepsin - Choris
737	22	78.6	243	2	S75630	hypothetical prote	810	22	78.6	324	2	S62735	hypothetical prote
738	22	78.6	243	2	C82393	probable amino aci	811	22	78.6	325	2	C89879	hypothetical prote
739	22	78.6	244	1	UESY25	vegetative storage	812	22	78.6	325	2	T24762	conserved hypothet
740	22	78.6	244	2	G75259	conserved hypothet	813	22	78.6	325	2	T05303	hypothetical prote
741	22	78.6	244	2	A83236	hypothetical prote	814	22	78.6	326	2	C33788	hypothetical prote
742	22	78.6	247	1	TRDG	trypsin (EC 3.4.21	815	22	78.6	327	2	AC1210	thiamin biosynthes
743	22	78.6	247	2	S12764	trypsin (EC 3.4.21	816	22	78.6	328	2	T33788	hypothetical prote
744	22	78.6	249	1	BVSCM5	transcription regu	817	22	78.6	329	2	B83985	branched-chain alp
745	22	78.6	250	2	D75054	hypothetical prote	818	22	78.6	330	2	B83985	branched-chain alp
746	22	78.6	250	2	B71107	hypothetical prote	819	22	78.6	331	2	AD1246	hypothetical prote
747	22	78.6	251	2	AI1927	ABC transporter AT	820	22	78.6	331	2	AH1608	branched-chain ami
748	22	78.6	252	2	G72618	hypothetical prote	821	22	78.6	331	2	C72239	branched-chain ami
749	22	78.6	253	2	T45446	probable two-compo	822	22	78.6	332	2	D81211	probable branched-
750	22	78.6	253	2	H69046	hypothetical prote	823	22	78.6	332	2	D81787	hypothetical prote
751	22	78.6	254	2	S08511	vegetative storage	824	22	78.6	332	2	AI2110	hypothetical prote
752	22	78.6	254	2	B73216	conserved hypothet	825	22	78.6	334	2	G93533	hypothetical prote
753	22	78.6	258	2	D83190	UDP-N-acetylglucos	826	22	78.6	334	2	G75344	probable polyferre
754	22	78.6	258	2	G84279	stationary phase s	827	22	78.6	336	2	H69105	ADP-glucose 4,6-d
755	22	78.6	259	2	D70342	hypothetical prote	828	22	78.6	338	2	H82615	3-oxoacyl-[ACP] sy
756	22	78.6	261	2	A70002	protein kinase hom	829	22	78.6	338	2	A84712	probable protein k
757	22	78.6	263	2	C83959	ribonuclease H rnh	830	22	78.6	338	2	D70382	phenylalanine-tRNA
758	22	78.6	264	2	H75406	hydrolase, alpha/b	831	22	78.6	339	2	G82113	lipoprotein-34 Nlp
759	22	78.6	264	2	S67861	GumL protein - Xan	832	22	78.6	339	2	G82113	

833	22	78.6	339	2	F84619	probable protein k	906	22	78.6	396	2	T40010	hypothetical prote
834	22	78.6	340	2	S74768	UDPglucose 4-epime	907	22	78.6	397	2	AE1142	penicillin-binding
835	22	78.6	341	2	A81394	galactosyltransfer	908	22	78.6	398	2	G69496	ATP-dependent 26S
836	22	78.6	342	2	A81769	galactosyltransfer	909	22	78.6	398	2	A58922	esterase/N-deacety
837	22	78.6	342	2	S62711	glutamate-ammonia	910	22	78.6	398	2	H71288	conserved hypothet
838	22	78.6	342	2	G75163	hypothetical prote	911	22	78.6	398	2	AH1500	penicillin-binding
839	22	78.6	342	2	D71178	hypothetical prote	912	22	78.6	398	2	AB2691	conserved hypothet
840	22	78.6	344	2	AD1200	conserved hypothet	913	22	78.6	400	1	OKBO2R	protein kinase (EC
841	22	78.6	345	2	T15599	hypothetical prote	914	22	78.6	400	2	AD2322	phosphoglycerate k
842	22	78.6	345	2	G69450	spermidine/putresc	915	22	78.6	401	1	OKMS2R	protein kinase (EC
843	22	78.6	346	1	F69590	aspartate-semialde	916	22	78.6	404	1	OKHU2R	protein kinase (EC
844	22	78.6	346	2	H83469	acetylpolypamine am	917	22	78.6	404	2	S64944	hypothetical prote
845	22	78.6	350	2	S71923	cysteine proteinas	918	22	78.6	408	2	S33683	site-specific DNA-
846	22	78.6	351	2	H69053	precorrin-3 methyl	919	22	78.6	410	2	T46034	branched-chain-ami
847	22	78.6	351	2	T37498	probable alpha-tub	920	22	78.6	410	2	T47586	hypothetical prote
848	22	78.6	352	2	A75081	aspartate kinase r	921	22	78.6	410	2	T50718	hypothetical prote
849	22	78.6	352	2	A69743	ATP-binding Mrp-li	922	22	78.6	410	2	H86290	hypothetical prote
850	22	78.6	353	2	T04567	O-sialoglycoprotei	923	22	78.6	411	2	G89074	protein K04A8.5 [i
851	22	78.6	356	2	C57010	3-dehydroquinat s	924	22	78.6	414	2	A75505	probable glutaryl-
852	22	78.6	356	2	A25918	thrombomodulin - b	925	22	78.6	414	2	T12022	3-oxoacyl-[acyl-ca
853	22	78.6	357	2	S72214	hypothetical prote	926	22	78.6	416	2	T45284	glutamate dehydrog
854	22	78.6	357	2	S76692	hypothetical prote	927	22	78.6	416	2	G72305	glutamate dehydrog
855	22	78.6	359	1	B69266	acetylpolypamine am	928	22	78.6	416	2	AD3075	sarcosine oxidase
856	22	78.6	359	2	S40746	hypothetical prote	929	22	78.6	416	2	E98211	sarcosine oxidase
857	22	78.6	361	2	S68268	apurinic/pyrimidi	930	22	78.6	416	2	AH2387	hypothetical prote
858	22	78.6	361	2	C72560	hypothetical prote	931	22	78.6	417	2	A45794	tubulin alpha chai
859	22	78.6	362	2	T29053	branched-chain-ami	932	22	78.6	420	2	T02132	probable serine/th
860	22	78.6	362	2	A47411	ADPribosylarginine	933	22	78.6	421	2	S38110	O-sialoglycoprotei
861	22	78.6	362	2	A38135	ADP-ribosylarginin	934	22	78.6	421	2	S38110	I7 protein - vacci
862	22	78.6	365	1	W8LHLS	E2 protein - human	935	22	78.6	423	1	W2V217	glutamate 5-kinase
863	22	78.6	366	2	B84712	probable protein k	936	22	78.6	423	2	S46741	topoisomerase II -
864	22	78.6	367	2	H70519	probable aceAa pro	937	22	78.6	423	2	C42511	I7L protein - vacc
865	22	78.6	369	2	AF0162	putrescine-binding	938	22	78.6	423	2	E36843	K7L protein - vari
866	22	78.6	370	2	AC2845	oxidoreductase Atu	939	22	78.6	423	2	T28499	hypothetical prote
867	22	78.6	370	2	E37622	gtn reductase (Y13	940	22	78.6	423	2	C72158	I7L protein - vari
868	22	78.6	372	1	W8L31	E2 protein - human	941	22	78.6	424	2	E75102	ribulose-bisphosph
869	22	78.6	372	2	H84044	citrate synthase I	942	22	78.6	425	2	D97937	hypothetical prote
870	22	78.6	373	2	A49806	prv43 protein - su	943	22	78.6	428	2	E95906	hypothetical prote
871	22	78.6	373	2	T45747	GDP-D-mannose-4,6-	944	22	78.6	430	2	AB2646	flagellin solute-bi
872	22	78.6	373	2	T45043	hypothetical prote	945	22	78.6	430	2	F71084	ribulose-bisphosph
873	22	78.6	374	2	S69339	Ig heavy chain V r	946	22	78.6	432	2	T36300	probable secreted
874	22	78.6	374	2	AH1825	protoporphyrin IX	947	22	78.6	432	2	T34154	hypothetical prote
875	22	78.6	377	2	B90432	citrate synthase [	948	22	78.6	435	2	H69607	alpha-ketoglutarat
876	22	78.6	378	2	T30965	hypothetical prote	949	22	78.6	436	2	A97428	flad protein (U951
877	22	78.6	378	2	AE2942	conserved hypothet	950	22	78.6	437	2	A75194	glycogen synthase
878	22	78.6	378	2	E98340	hypothetical prote	951	22	78.6	438	2	D71924	transcription term
879	22	78.6	379	2	T44615	citrate (si)-synth	952	22	78.6	438	2	F64588	transcription term
880	22	78.6	379	2	T19773	hypothetical prote	953	22	78.6	439	2	AE1251	probable peptidogl
881	22	78.6	379	2	T49919	hypothetical prote	954	22	78.6	443	2	G64114	site-specific DNA-
882	22	78.6	382	2	D82516	type I restriction	955	22	78.6	445	2	T05887	hypothetical prote
883	22	78.6	382	2	F81298	probable decarboxy	956	22	78.6	448	2	H95151	NADP-specific glut
884	22	78.6	383	2	B64205	S-adenosylmethioni	957	22	78.6	448	2	D98019	glutamate dehydrog
885	22	78.6	383	2	F87719	protein Kil9.2 [im	958	22	78.6	448	2	E96991	Na+/H+ antiporter,
886	22	78.6	385	2	S54987	reverse transcript	959	22	78.6	449	2	S13336	tubulin alpha-1 ch
887	22	78.6	385	2	S54988	reverse transcript	960	22	78.6	449	2	A54760	translation elonga
888	22	78.6	385	2	S54990	reverse transcript	961	22	78.6	449	2	JC5117	translation elonga
889	22	78.6	385	2	S54991	reverse transcript	962	22	78.6	451	2	S45051	tubulin alpha-B ch
890	22	78.6	385	2	S54992	reverse transcript	963	22	78.6	452	2	A57217	tipE protein - fru
891	22	78.6	385	2	S54995	reverse transcript	964	22	78.6	452	2	G72770	hypothetical prote
892	22	78.6	385	2	S54997	reverse transcript	965	22	78.6	453	2	AE0774	probable protease
893	22	78.6	387	2	F95069	transmembrane prot	966	22	78.6	453	2	T29103	hypothetical prote
894	22	78.6	387	2	H82176	carboxynorspermidi	967	22	78.6	454	2	AF0896	L-serine ammonia-1
895	22	78.6	387	2	G71097	probable amidohydr	968	22	78.6	454	2	H91127	L-serine deaminase
896	22	78.6	388	2	B82317	hypothetical prote	969	22	78.6	456	2	G85972	probable L-serine
897	22	78.6	389	1	T48230	serine/threonine-s	970	22	78.6	456	2	H97076	S-adenosylmethioni
898	22	78.6	389	1	J80443	alanine racemase, (	971	22	78.6	456	2	F97810	magnesium transpor
899	22	78.6	389	2	D90498	acyl-CoA dehydroge	972	22	78.6	456	2	C71663	magnesium transpor
900	22	78.6	389	2	F97692	glutamate 5-kinase	973	22	78.6	457	2	D96735	hypothetical prote
901	22	78.6	389	2	AC2918	glutamate 5-kinase	974	22	78.6	459	2	D36889	3-isopropylmalate
902	22	78.6	393	2	SG1684	glutamine	975	22	78.6	460	2	S35134	probable 3-isoprop
903	22	78.6	394	2	JC4039	actin-like protein	976	22	78.6	460	2	D86777	hypothetical prote
904	22	78.6	394	2	E75104	hypothetical prote	977	22	78.6	461	2	B89972	fumarate hydratase
905	22	78.6	395	2	S74051	hypothetical prote	978	22	78.6	461	2	D89894	hypothetical prote

979 22 78.6 461 2 D86294 T24D19.8 protein -  
 980 22 78.6 462 2 A51323 3-isopropylmalate  
 981 22 78.6 462 2 AF1694 3-isopropylmalate  
 982 22 78.6 464 2 AG0347 probable proteinase  
 983 22 78.6 465 1 T51095 acid phosphatase (  
 984 22 78.6 465 2 T51094 acid phosphatase (  
 985 22 78.6 466 2 T45269 fumarate hydratase  
 986 22 78.6 466 2 T43727 fumarate hydratase  
 987 22 78.6 466 2 E95262 L-seryl-tRNA<sup>Sec</sup> se  
 988 22 78.6 468 1 B84540 acid phosphatase (  
 989 22 78.6 470 1 T51096 acid phosphatase (  
 990 22 78.6 471 2 S05392 hypothetical prote  
 991 22 78.6 474 1 JG6042 cytosol nonspecifi  
 992 22 78.6 474 1 H82564 3-isopropylmalate  
 993 22 78.6 475 1 YWBO tryptophan-tRNA li  
 994 22 78.6 476 2 T47278 cysteine-tRNA liga  
 995 22 78.6 476 2 G84634 probable polylcar  
 996 22 78.6 477 2 H85981 D-alanyl-D-alanine  
 997 22 78.6 477 2 E91136 D-alanyl-D-alanine  
 998 22 78.6 477 2 A54535 serine-type D-Ala-  
 999 22 78.6 478 2 H69075 nitrogenase iron-m  
 1000 22 78.6 479 2 B89764 transcription regu

## ALIGNMENTS

RESULT 1  
 JU0240  
 nuclear matrix protein N/MAX-74 - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Jul-2004  
 C:Accession: JU0240  
 R:Ingaki, H.; Matsushima, Y.; Ohoshima, M.; Nakamura, K.; Kadowaki, T.; Kitagawa, Y.  
 submitted to JIPID, October 1993  
 A:Description: Twice repeated amino acid sequence segments common to Matrin 3 and a nov  
 A:Reference number: JU0240  
 A:Accession: JU0240  
 A:Molecule type: protein  
 A:Residues: 1-74 <INA>  
 A:Cross-references: UNIPROT:Q14966  
 A:Experimental source: Namalwa cell; Hela cell  
 C:Keywords: nuclear matrix

Query Match 96.4%; Score 27; DB 2; Length 74;  
 Best Local Similarity 83.3%; Pred. No. 15;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 18 GYSVEE 23

RESULT 2  
 B87323  
 chemotaxis protein CheYv [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 16-Aug-2004  
 C:Accession: B87323  
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; DeBoY, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 n. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: B87249; MUID:21173698; PMID:11259647

A:Accession: B87323  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-122 <STO>  
 A:Cross-references: UNIPROT:Q9AAK1; GB:AE005673; NID:gl13421798; PIDN:AAK22582.1; GSPDB:C  
 C:Genetics:  
 A:Gene: CC0596  
 C:Superfamily: Signal transduction receiver (phosphoacceptor) protein, CheY type; respon

Query Match 96.4%; Score 27; DB 2; Length 122;  
 Best Local Similarity 83.3%; Pred. No. 24;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 25 GYAVEE 30

RESULT 3  
 G75075  
 hypothetical protein PAB1706 - Pyrococcus abyssi (strain Orsay)  
 C:Species: Pyrococcus abyssi  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C:Accession: G75075  
 R:anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc  
 A:Reference number: A75001  
 A:Accession: G75075  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-233 <KAW>  
 A:Cross-references: UNIPROT:Q9UZZ8; GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB49906  
 A:Experimental source: strain Orsay  
 C:Genetics:  
 A:Gene: PAB1706  
 C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1634

Query Match 96.4%; Score 27; DB 2; Length 233;  
 Best Local Similarity 83.3%; Pred. No. 48;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 18 GYSVEE 23

RESULT 4  
 WMBP8H  
 gene 8.5 protein - phage PZA  
 C:Species: phage PZA  
 A:Note: host Bacillus subtilis  
 C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 09-Jul-2004  
 C:Accession: C24831  
 R:Paces, V.; Vlcek, C.; Urbanek, P.  
 Gene 44, 107-114, 1986  
 A:Title: Nucleotide sequence of the late region of Bacillus subtilis phage PZA, a close  
 A:Reference number: A91550; MUID:87031573; PMID:3095188  
 A:Accession: C24831  
 A:Molecule type: DNA  
 A:Residues: 1-280 <PAC>  
 A:Cross-references: UNIPROT:P07532; GB:M1813; GB:M13904; GB:M13905; NID:g216046; PIDN:AF  
 A:Note: the authors translated the codon CTT for residue 207 as Pro, GGG for residue 208  
 2 as Val, and GCG for residue 213 as Ser  
 C:Genetics:  
 A:Gene: 8.5  
 C:Superfamily: phage Nf assembly protein  
 C:Keywords: capsid assembly; late protein

Query Match 96.4%; Score 27; DB 1; Length 280;  
 Best Local Similarity 83.3%; Pred. No. 58;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 96 GYAVEE 101

RESULT 5  
 WMBP8F  
 gene 8.5 protein - phage phi-29

N;Alternate names: assembly protein; head fiber protein  
 C;Species: phage phi-29  
 A;Note: host *Bacillus subtilis*  
 C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 09-Jul-2004  
 C;Accession: C25816  
 R;VLcek, C.; Paces, V.  
 Gene 46, 215-225, 1986  
 A;Title: Nucleotide sequence of the late region of *Bacillus* phage phi-29 completes the 1  
 A;Reference number: A25816; MUID:87106857; PMID:3803926  
 A;Accession: C25816  
 A;Molecule type: DNA  
 A;Residues: 1-280 <VLC>  
 A;Cross-references: UNIPROT:P20344; GB:M14782; NID:G215323; PIDN:AAA322281.1; PID:G215326  
 C;Genetics:  
 C;Superfamily: phage Nf assembly protein  
 C;Keywords: capsid assembly; head protein; late protein

Query Match 96.4%; Score 27; DB 1; Length 280;  
 Best Local Similarity 83.3%; Pred. No. 58;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 96 GYAVEE 101

RESULT 6  
 A70441  
 Glycerol-3-phosphate dehydrogenase (NAD) (EC 1.1.1.8) - Aquifex aeolicus  
 C;Species: Aquifex aeolicus  
 C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 16-Aug-2004  
 C;Accession: A70441  
 R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.  
 Nature 392, 353-358, 1998  
 A;Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
 A;Reference number: A70300; MUID:98196666; PMID:9537320  
 A;Accession: A70441  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-313 <AQP>  
 A;Cross-references: GB:AE000748; NID:G2983960; PIDN:AAC07511.1; PID:G2983964; GB:AE00065  
 A;Experimental source: strain VP5  
 C;Genetics:  
 A;Gene: gspA  
 C;Superfamily: Glycerol-3-phosphate dehydrogenase (NAD)  
 C;Keywords: oxidoreductase

Query Match 96.4%; Score 27; DB 2; Length 313;  
 Best Local Similarity 83.3%; Pred. No. 65;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 246 GYSVEE 251

RESULT 7  
 YFBSA  
 phenylalanine-tRNA ligase (EC 6.1.1.20) alpha chain - *Bacillus subtilis*  
 N;Alternate names: phenylalanyl-tRNA synthetase alpha chain  
 C;Species: *Bacillus subtilis*  
 C;Date: 30-Sep-1992 #sequence\_revision 02-Jul-1998 #text\_change 09-Jul-2004  
 C;Accession: H69675; I40459; S11730  
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertez  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
 A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrazi, E.  
 Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler  
 ich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardiniois,  
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon,  
 A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,  
 teuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.  
 A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377  
 A;Accession: H69675  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-344 <KUN>  
 A;Cross-references: UNIPROT:P17921; GB:Z99118; GB:AL009126; NID:G2635200; PIDN:CAB14824.1  
 A;Experimental source: strain 168  
 R;Brakhage, A.A.; Wozny, M.; Putzer, H.  
 Biochimie 72, 725-734, 1990  
 A;Title: Structure and nucleotide sequence of the *Bacillus subtilis* phenylalanyl-tRNA synthetase  
 A;Reference number: I40459; MUID:91175935; PMID:2127701  
 A;Accession: I40459  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-89, 'DRQLTSCRPCCSR', 106-344 <BRA>  
 A;Cross-references: EMBL:X53057; NID:G40052; PIDN:CAA37224.1; PID:G40053  
 C;Genetics:  
 A;Gene: pheS  
 C;Complex: heterotetramer; two alpha and two beta chains; (see PIR:YFBSB)  
 C;Function:  
 A;Description: catalyzes transfer of activated phenylalanine to phenylalanyl-tRNA  
 C;Superfamily: phenylalanine-tRNA ligase alpha chain  
 C;Keywords: aminoacyl-tRNA synthetase; ATP; heterotetramer; ligase; protein biosynthesis

Query Match 96.4%; Score 27; DB 1; Length 344;  
 Best Local Similarity 83.3%; Pred. No. 72;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 125 GYTVEE 130

RESULT 8  
 B83605  
 probable acetylpolymine aminohydrolase PA0321 [imported] - *Pseudomonas aeruginosa* (strain  
 C;Species: *Pseudomonas aeruginosa*  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C;Accession: B83605  
 R;Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
 A;Reference number: A82950; MUID:20437337; PMID:10984043  
 A;Accession: B83605  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-344 <STO>  
 A;Cross-references: UNIPROT:Q916H0; GB:AE004470; NID:G9946164; PIDN:AAG03711  
 A;Experimental source: strain PA01  
 C;Genetics:  
 A;Gene: PA0321

Query Match 96.4%; Score 27; DB 2; Length 344;  
 Best Local Similarity 83.3%; Pred. No. 72;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 322 GYAVEE 327

RESULT 9  
 A40990  
 GTP-binding regulatory protein G-alpha-4 - slime mold (*Dictyostelium discoideum*) (fragment

C:\Genetic8:



Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||

Db 243 GYSVEE 248

RESULT 14

A57506  
palindrome-binding protein - fungus (*Fusarium solani*)  
C:Species: *Fusarium solani* f. sp. *psii*  
C:Date: 08-Dec-1995 #sequence\_revision 08-Dec-1995 #text\_change 09-Jul-2004  
C:Accession: A57506  
R:Li, D.; Kolattukudy, P.E.  
J. Biol. Chem. 270, 11753-11756, 1995  
A:Title: Cloning and expression of cDNA encoding a protein that binds a palindromic promoter  
A:Reference number: A57506; MUID:95263512; PMID:7744822  
A:Accession: A57506  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-457 <LTA>  
A:Cross-references: UNIPROT:Q00858; GB:U23722; NID:g763041; PID:g763042  
C:Superfamily: Neurospora crassa zinc finger protein wc-2  
C:Keywords: DNA binding

Query Match 96.4%; Score 27; DB 2; Length 457;  
Best Local Similarity 83.3%; Pred. No. 97;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||

Db 139 GYSVEE 144

RESULT 15

C86739  
hypotheical protein tagH [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)  
C:Species: *Lactococcus lactis* subsp. *lactis*  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: C86739  
R:Boletín, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp.  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: C86739  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-466 <STO>  
A:Cross-references: UNIPROT:Q9CH26; GB:AE005176; PID:g12723849; PIDN:AAK05013.1; GSPDB:G  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: tagH  
C:Superfamily: *Bacillus subtilis* teichoic acid ABC transporter tagH; ATP-binding cassette

Query Match 96.4%; Score 27; DB 2; Length 466;  
Best Local Similarity 83.3%; Pred. No. 99;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||

Db 421 GYSVEE 426

RESULT 16

F90032  
urocanate hydratase [imported] - *Staphylococcus aureus* (strain N315)  
C:Species: *Staphylococcus aureus*  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: F90032  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: F90032  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-553 <KUR>  
A:Cross-references: UNIPROT:Q99RU2; GB:BA000018; PID:g13702131; PIDN:BA843423.1; GSPDB:G  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: hutU  
C:Superfamily: urocanate hydratase

Query Match 96.4%; Score 27; DB 2; Length 553;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||

Db 270 GYSVEE 275

RESULT 17

AB2168  
two-component hybrid sensor and regulator all2897 [imported] - *Nostoc* sp. (strain PCC 712;  
C:Species: *Nostoc* sp. PCC 7120  
A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: AB2168  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium* Anal  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AB2168  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-646 <KUR>  
A:Cross-references: UNIPROT:Q8VT29; GB:BA000019; PIDN:BA874596.1; PID:g17131991; GSPDB:G  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all2897

Query Match 96.4%; Score 27; DB 2; Length 646;  
Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||

Db 314 GYSVEE 319

RESULT 18

A44919  
GCR3 protein - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: protein YMS64.07; protein YMS53.01; protein YMR125w  
C:Species: *Saccharomyces cerevisiae*  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C:Accession: A44919; S53055; S54494  
R:Uemura, H.; Jigami, Y.  
J. Bacteriol. 174, 5526-5532, 1992  
A:Title: GCR3 encodes an acidic protein that is required for expression of glycolytic ge  
A:Reference number: A44919; MUID:92380925; PMID:1512188  
A:Accession: A44919  
A:Molecule type: DNA  
A:Residues: 1-858 <UEM>  
A:Cross-references: UNIPROT:P34160; GB:D10224; NID:g464221; PIDN:BA801076.1; PID:d100154;  
A:Note: sequence extracted from NCBI backbone (NCBIN:112104, NCBI:P:112106)  
R:Badcock, K.; Churcher, C.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S53055  
A:Accession: S53055  
A:Molecule type: DNA  
A:Residues: 339-858 <BAD>

A;Cross-references: EMBL:Z48622; NID:g728663; PIDN:CAA88550.1; PID:g728664; MIPS:YMR125W  
R;Lye, G.; Churcher, C.M.  
submitted to the EMBL Data Library, May 1995  
A;Reference number: S54014  
A;Accession: S54494  
A;Molecule type: DNA  
A;Residues: 'MFNRKRG', 6-489 <LYE>  
A;Cross-references: EMBL:Z49273; NID:g809577; PIDN:CAA89274.1; PID:g809584; MIPS:YMR125W  
C;Genetics:  
A;Gene: SGD:STO1; GCR3  
A;Cross-references: MIPS:YMR125W; SGD:S0004732  
A;Map position: 13R  
C;Keywords: DNA binding; nucleus

Query Match 96.4%; Score 27; DB 2; Length 858;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 565 GYTVEE 570

RESULT 19  
AC2542  
cation-transporting ATPase alr7622 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AC2542  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AC2542  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-879 <KUR>  
A;Cross-references: UNIPROT:Q82590; GB:AP003602; PIDN:BA877265.1; PID:gl7134707; GSPDB:G  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: alr7622  
A;Genome: plasmid  
C;Superfamily: Bacillus probable copper-transporting ATPase yvgX; ATPase nucleotide-bind

Query Match 96.4%; Score 27; DB 2; Length 879;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 200 GYTVEE 205

RESULT 20  
A84888  
hypothetical protein At2g45230 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: A84888  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: A84888  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1374 <STO>  
A;Cross-references: UNIPROT:O22148; GB:AE002093; NID:g2583130; PIDN:AAB82639.1; GSPDB:GN  
C;Genetics:

A;Gene: At2g45230  
A;Map position: 2

Query Match 96.4%; Score 27; DB 2; Length 1374;  
Best Local Similarity 83.3%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 403 GYTVEE 408

RESULT 21  
C75026  
hypothetical protein PAB3433 - Pyrococcus abyssi (strain Orsay)  
C;Species: Pyrococcus abyssi  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: C75026  
R;anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc  
A;Reference number: A75001  
A;Accession: C75026  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-80 <KAM>  
A;Cross-references: UNIPROT:Q9UXM4; GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAB5064;  
A;Experimental source: strain Orsay  
C;Genetics:  
A;Gene: PAB3433  
C;Superfamily: conserved hypothetical protein MJ1593

Query Match 92.9%; Score 26; DB 2; Length 80;  
Best Local Similarity 83.3%; Pred. No. 28;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 31 GYKVEE 36

RESULT 22  
B64631  
hypothetical protein HP0893 - Helicobacter pylori (strain 26695)  
C;Species: Helicobacter pylori  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C;Accession: B64631  
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N  
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: B64631  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-95 <TOM>  
A;Cross-references: UNIPROT:O25553; GB:AE000599; GB:AE000511; NID:g2314028; PIDN:AA07094

Query Match 92.9%; Score 26; DB 2; Length 95;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 60 GYTVEE 65

RESULT 23  
C71883  
hypothetical protein jhp0832 - Helicobacter pylori (strain J99)  
C;Species: Helicobacter pylori

A;Variety: strain J99  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: C71883  
R;Alt: R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
Nature 397, 176-180, 1999  
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A;Reference number: A71800; MUID:99120557; PMID:9923682  
A;Accession: C71883  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-95 <ARN>  
A;Cross-references: UNIPROT:Q9ZKV3; GB:AE001512; GB:AE001439; NID:g4155392; PIDN:AAD0641  
A;Experimental source: strain J99  
C;Genetics:  
A;Gene: jhp0832

Query Match 92.9%; Score 26; DB 2; Length 95;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 60 GYTIIE 65

RESULT 24  
AE1087  
hypotheical protein lmo0100 [imported] - Listeria monocytogenes (strain EGD-e)  
C;Species: Listeria monocytogenes  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AE1087  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A;Title: Comparative genomics of Listeria species.  
A;Accession number: AB1077; MUID:21537279; PMID:11679669  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-116 <GUA>  
A;Cross-references: UNIPROT:Q8VAL8; GB:NC\_003210; PIDN:CAC98315.1; PID:g16409459; GSPDB:  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmo0100

Query Match 92.9%; Score 26; DB 2; Length 116;  
Best Local Similarity 83.3%; Pred. No. 41;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 80 GYRVEE 85

RESULT 25  
AD1451  
hypotheical protein lin0147 [imported] - Listeria innocua (strain Clip11262)  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AD1451  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AD1451

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-116 <GUA>  
A;Cross-references: UNIPROT:Q92FF9; GB:AL592022; PIDN:CAC95380.1; PID:g16412566; GSPDB:G  
C;Genetics:  
A;Gene: lin0147

Query Match 92.9%; Score 26; DB 2; Length 116;  
Best Local Similarity 83.3%; Pred. No. 41;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 80 GYRVEE 85

RESULT 26  
C82805  
DNA-binding protein XF0446 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: C82805  
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717; PMID:10910347  
A;Note: for a complete list of authors see reference number A59328 below  
A;Accession: C82805  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-124 <SIM>  
A;Cross-references: UNIPROT:Q9PG56; GB:AE003895; GB:AE003849; NID:g9105283; PIDN:AAF8325  
A;Experimental source: strain 9a5c  
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Pacincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.  
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A;Reference number: A59328  
A;Contents: annotation  
C;Genetics:  
A;Gene: XF0446

Query Match 92.9%; Score 26; DB 2; Length 124;  
Best Local Similarity 66.7%; Pred. No. 44;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 48 GYTIIE 53

RESULT 27  
T45079  
hypotheical protein frxA [imported] - Pyrococcus furiosus  
C;Species: Pyrococcus furiosus  
C;Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 09-Jul-2004  
C;Accession: T45079  
R;Kletzin, A.; Adams, M.W.  
J. Bacteriol. 178, 248-257, 1996  
A;Title: Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate fr  
ima.  
A;Reference number: Z22908; MUID:96125254; PMID:8550425  
A;Accession: T45079  
A;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-137 <KLE>  
A:Cross-references: UNIPROT:Q51795; EMBL:X85250; NID:g1197352; PIDN:CAA59496.1; PID:g119  
A:Experimental source: DSM 3638  
C:Genetics:  
A:Gene: frxA  
C:Superfamily: [NiFe]-hydrogenase maturation protease

Query Match 92.9%; Score 26; DB 2; Length 137;  
Best Local Similarity 83.3%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 18 GYKVEE 23

RESULT 28  
C71113  
Ni,Fe-Hydrogenase maturation protease (EC 3.4.24.-) - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 16-Aug-2004  
C:Accession: C71113  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A:Reference number: A71000; MUID:98344137; PMID:9679194  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Accession: C71113  
A:Molecule type: DNA  
A:Residues: 1-151 <KAW>  
A:Cross-references: GB:AF000003; NID:g3236130; PIDN:BA29765.1; PID:g3257082  
A:Experimental source: strain OF3  
A>Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0674  
C:Superfamily: [NiFe]-hydrogenase maturation protease  
C:Keywords: hydrolase; metalloproteinase

Query Match 92.9%; Score 26; DB 2; Length 151;  
Best Local Similarity 83.3%; Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 30 GYRVEE 35

RESULT 29  
AG1132  
Hypothetical protein lmo0462 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AG1132  
R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A>Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AG1132  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-161 <GLA>  
A:Cross-references: UNIPROT:Q8Y9R1; GB:NC\_003210; PIDN:CAC98541.1; PID:g16409839; GSPDB:  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0462

Query Match 92.9%; Score 26; DB 2; Length 161;

Best Local Similarity 66.7%; Pred. No. 58;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 125 GYAIEE 130

RESULT 30  
T01785  
protoporphyrin IX magnesium chelatase (EC 4.99.1.-) - common tobacco (fragment)  
C:Species: Nicotiana tabacum (common tobacco)  
C>Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T01785  
R:Kruse, E.; Mock, H.P.; Grimm, B.  
submitted to the EMBL Data Library, July 1997  
A:Description: Isolation and characterisation of tobacco (Nicotiana tabacum) cDNA clones  
A:Reference number: Z14428  
A:Accession: T01785  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-177 <KRU>  
A:Cross-references: UNIPROT:O22434; EMBL:AF014051; NID:g2338715; PIDN:AAB67235.1; PID:g2;  
A:Experimental source: strain SR 1  
C:Genetics:  
A:Gene: ChlH  
C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase  
C:Keywords: chlorophyll biosynthesis; lyase; magnesium

Query Match 92.9%; Score 26; DB 2; Length 177;  
Best Local Similarity 83.3%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 114 GYLVEE 119

RESULT 31  
T13309  
Hypothetical protein 20 - Streptococcus phage phi-O1205  
C:Species: Streptococcus phage phi-O1205  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: T13309  
R:Stanley, E.; Fitzgerald, G.F.; Le Marrec, C.; Fayard, B.; van Sinderen, D.  
Microbiology 143, 3417-3429, 1997  
A>Title: Sequence analysis and characterization of phi O1205, a temperate bacteriophage i  
A:Reference number: Z17654; MUID:98048466; PMID:9387220  
A:Accession: T13309  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-181 <STA>  
A:Cross-references: UNIPROT:O34051; EMBL:U88974; NID:g2444080; PID:g2454217; PIDN:AAC795;  
A:Experimental source: host Streptococcus thermophilus strain CNRZ1205  
C:Superfamily: Streptococcus phage phi-O1205 hypothetical protein 20

Query Match 92.9%; Score 26; DB 2; Length 181;  
Best Local Similarity 83.3%; Pred. No. 65;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 148 GYDVEE 153

RESULT 32  
H72256  
conserved hypothetical protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: H72256  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.  
 Nature 399, 323-329, 1999  
 A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing  
 A;Reference number: A72200; MUID:99287316; PMID:10360571  
 A;Accession: H72256  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-204 <ARN>  
 A;Cross-references: UNIPROT:Q9X1E1; GB:AE001794; GB:AE000512; NID:g4981978; PIDN:AAD3649  
 A;Experimental source: strain MSB8  
 C;Genetics:  
 A;Gene: TM1427

Query Match	92.9%;	Score 26;	DB 2;	Length 204;
Best Local Similarity	83.3%;	Pred. No. 74;		
Matches	5;	Conservative 0;	Mismatches 1;	Indels 0;
Gaps	0;			

QY 1 GYXVEE 6  
 |||||  
 DB 62 GYDVEE 67

RESULT 33  
 S39103  
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b (cob) - Sanduri wheat mitochondrion  
 C;Species: mitochondrion Triticum timopheevi (Sanduri wheat)  
 C;Date: 31-Dec-1993 #sequence\_revision 24-Feb-1995 #text\_change 09-Jul-2004  
 C;Accession: S39103; S42740  
 R;Zanlungo, S.; Begu, D.; Quinones, V.; Araya, A.; Jordana, X.  
 Curr. Genet. 24, 344-348, 1993  
 A;Title: RNA editing of apocytochrome b (cob) transcripts in mitochondria from two genera  
 A;Reference number: S36919; MUID:94073991; PMID:7504589  
 A;Accession: S39103  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-207 <ZAN>  
 A;Cross-references: UNIPROT:Q7M2D4  
 R;Saalaoui, E.; Litvak, S.; Araya, A.  
 Plant Sci. 66, 237-246, 1990  
 A;Title: The apocytochrome b from an alloplasmic line of wheat (T. aestivum, cytoplasm-T)  
 A;Reference number: S42740  
 A;Accession: S42740  
 A;Molecule type: DNA  
 A;Residues: 168-183,'H',185-207 <SAA>  
 A;Note: the amino acid sequence from fig.4 is inconsistent with the nucleotide sequence  
 C;Genetics:  
 A;Gene: cob; cob  
 A;Genome: mitochondrion  
 C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol  
 C;Keywords: electron transfer; heme; mitochondrion; oxidative phosphorylation; oxidoreductase  
 P;12-123/Domain: cytochrome b homology (fragments) <CBH>  
 P;12-76/Domain: cytochrome b6 homology (fragments) <CB6>  
 P;87-123/Domain: plastoquinol-plastocyanin reductase 17K protein homology (fragments) <17K>

Query Match	92.9%;	Score 26;	DB 2;	Length 207;
Best Local Similarity	66.7%;	Pred. No. 75;		
Matches	4;	Conservative 1;	Mismatches 1;	Indels 0;
Gaps	0;			

QY 1 GYXVEE 6  
 |||||  
 DB 183 GYSIEE 188

RESULT 34  
 AH1180  
 amino acid ABC transporter, ATP-binding protein homolog lmo0848 [imported] - Listeria monocytogenes  
 C;Species: Listeria monocytogenes  
 C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C;Accession: AH1180  
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.;Title: Comparative genomics of Listeria species.  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AH1180  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-242 <GLA>  
 A;Cross-references: UNIPROT:Q8Y8P8; GB:NC\_003210; PIDN:CAC98926.1; PID:g16410236; GSPDB:A  
 A;Experimental source: strain EGD-e  
 C;Genetics:  
 A;Gene: lmo0848  
 C;Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology

Query Match	92.9%;	Score 26;	DB 2;	Length 242;
Best Local Similarity	83.3%;	Pred. No. 88;		
Matches	5;	Conservative 0;	Mismatches 1;	Indels 0;
Gaps	0;			

QY 1 GYXVEE 6  
 |||||  
 DB 214 GYIVEE 219

RESULT 35  
 AI1537  
 amino acid ABC transporter, ATP-binding protein homolog lin0841 [imported] - Listeria innocua  
 C;Species: Listeria innocua  
 C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C;Accession: AI1537  
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.;Title: Comparative genomics of Listeria species.  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AI1537  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-242 <GLA>  
 A;Cross-references: UNIPROT:Q92DH2; GB:AL592022; PIDN:CAC96073.1; PID:g16413292; GSPDB:G  
 A;Experimental source: strain Clip11262  
 C;Genetics:  
 A;Gene: lin0841  
 C;Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology

Query Match	92.9%;	Score 26;	DB 2;	Length 242;
Best Local Similarity	83.3%;	Pred. No. 88;		
Matches	5;	Conservative 0;	Mismatches 1;	Indels 0;
Gaps	0;			

QY 1 GYXVEE 6  
 |||||  
 DB 214 GYIVEE 219

RESULT 36  
 NDECF5  
 type II site-specific deoxyribonuclease (EC 3.1.21.4) EcoRV - Escherichia coli  
 A;Alternate names: endonuclease EcoRV; type II restriction enzyme EcoRV  
 C;Species: Escherichia coli  
 C;Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 09-Jul-2004  
 C;Accession: A00784; I77184  
 R;Bougueleret, L.; Schwarzeisen, M.; Tsugita, A.; Zabeau, M.  
 Nucleic Acids Res. 12, 3659-3676, 1984  
 A;Title: Characterization of the genes coding for the Eco RV restriction and modification  
 A;Reference number: A93516; MUID:84221388; PMID:6328432  
 A;Accession: A00784  
 A;Molecule type: DNA  
 A;Residues: 1-245 <BOU>  
 A;Cross-references: UNIPROT:P04390; GB:X00530; GB:K02335; NID:g41324; PIDN:CAA25208.1; P  
 A;Note: the authors translated the codon GAT for residue 90 as Aen  
 R;Kraev, A.S.; Kravets, A.N.; Chernov, B.K.; Skryabin, K.G.; Baev, A.A.

```

Mol. Biol. 19, 236-242, 1985
A:Title: The EcoRV restriction-modification system: Genes, enzymes, synthetic substrates
A:Reference number: 157396
A:Accession: 177184
A:Molecule type: DNA
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-245 <RES>
A:Cross-references: GB:M19941; NID:g147784; PIDN:AAA24615.1; PID:g147786
C:Function:
A:Description: this is one of the type II restriction enzymes, which require only magnesium within, or at a short specific distance from, the recognition sites; the recognition of the recognition sequence
C:Superfamily: type II restriction endonuclease, EcoRV type
C:Keywords: endonuclease; hydrolase; magnesium

Query Match          92.9%; Score 26; DB 1; Length 245;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
        |||||
Db      60 GYVEE 65

RESULT 37
T06441
Storage protein homolog, 31K - soybean (fragment)
C:Species: Glycine max (soybean)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C:Accession: T06441
R:Mason, H.S.; Guerrero, F.D.; Boyer, J.S.; Mullet, J.E.
Plant Mol. Biol. 11, 845-856, 1988
A:Title: Proteins homologous to leaf glycoproteins are abundant in stems of darkgrown soybeans
A:Reference number: S08511
A:Accession: T06441
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-249 <MAS>
A:Cross-references: EMBL:M37529; NID:g169899; PID:g169900
C:Superfamily: vegetative storage protein; glucose-6-phosphatase catalytic domain homolog
C:Keywords: storage protein

Query Match          92.9%; Score 26; DB 2; Length 249;
Best Local Similarity 83.3%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
        |||||
Db      125 GYGVVE 130

RESULT 38
C03837
Hypothetical protein BH1499 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: C83837
R:Takami, H.; Nakasono, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Fujii, F.; Hiraoka, Y.
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and its relationship to the complete genome sequence of the alkaliphilic bacterium Halobacterium salinarum
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C83837
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-252 <STO>
A:Cross-references: UNIPROT:Q9KCS0; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA0522
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1499

Query Match          92.9%; Score 26; DB 2; Length 252;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GYXVEE 6
        |||||
Db      195 GYVEE 200

```

## RESULT 39

## UESY27

vegetative storage protein, 27K, precursor - soybean  
 N:Alternate names: 31K glycoprotein precursor  
 C:Species: Glycine max (soybean)  
 C:Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 09-Jul-2004  
 C:Accession: JN0697; JAO139; S08512; B45504; JAO158  
 R:Rapp, W.D.; Lilley, G.G.; Nielsen, N.C.  
 Theor. Appl. Genet. 79, 785-792, 1990  
 A:Title: Characterization of soybean vegetative storage proteins and genes.

```

A:Reference number: JN0697
A:Accession: JN0697
A:Molecule type: DNA
A:Residues: 1-254 <RAP>
A:Cross-references: UNIPROT:P10743
R:Staaswick, P.E.
Plant Physiol. 87, 250-254, 1988
A:Title: Soybean vegetative storage protein structure and gene expression.
A:Reference number: JAO139
A:Accession: JAO139
A:Molecule type: mRNA
A:Residues: 1-254 <STA>
A:Experimental source: strain Merr. cv Williams
A:Note: the cDNA clones corresponding to two different vegetative storage proteins were

```

```

A:Accession: S08512
A:Molecule type: mRNA
A:Residues: 1-254 <MAS>
A:Cross-references: EMBL:X17414; NID:g18761; PIDN:CAA35464.1; PID:g18762
R:Staaswick, P.E.
Plant Physiol. 89, 717, 1989
A:Reference number: A45504
A:Contents: erratum
A:Accession: B45504
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 207-254 <ST2>
C:Comment: Vegetative storage protein is found in vegetative tissues but not in seeds.
C:Genetics:
A:Gene: VSP27

```

```

C:Superfamily: vegetative storage protein; glucose-6-phosphatase catalytic domain homolog
C:Keywords: glycoprotein; storage protein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:36-254/Product: 27K vegetative storage protein #status predicted <MAP>
F:130/Binding site: carbonyl site: carbonyl site (Asn) (covalent) #status predicted

```

```

Query Match          92.9%; Score 26; DB 1; Length 254;
Best Local Similarity 83.3%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GYXVEE 6
        |||||
Db      123 GYGVVE 128

```

## RESULT 40

## H84713

Hypothetical protein At2g30890 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: H84713  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84713

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-257 <STO>

A:Cross-references: UNIPROT:O80854; GB:AE002093; NID:g3201616; PIDN:AAC20723.1; GSPDB:GN

C:Genetics:

A:Gene: At2G30890

A:Map position: 2

Query Match 92.9%; Score 26; DB 2; Length 257;  
Best Local Similarity 83.3%; Pred. No. 94;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 244 GYVEE 249

RESULT 41

S75190

Hypothetical protein elr2042 - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004

A:Accession: S75190

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996

A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75190

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-275 <KAN>

A:Cross-references: UNIPROT:P73079; EMBL:D90903; GB:AB001339; NID:g1652127; PIDN:BAAL710

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 92.9%; Score 26; DB 2; Length 275;  
Best Local Similarity 83.3%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 131 GYVEE 136

RESULT 42

H69802

conserved hypothetical protein yfiE - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

A:Accession: H69802

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallei iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsstra, P.; Tognoni, A.; Tosato, V.; Uchiyama T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: H69802

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-285 <KUN>

A:Cross-references: UNIPROT:P54721; GB:Z99108; GB:AL009126; NID:g2633055; PIDN:CAB12653.1

A:Experimental source: strain 168

C:Genetics:

A:Gene: yfiE

Query Match 92.9%; Score 26; DB 2; Length 285;

Best Local Similarity 66.7%; Pred. No. 1e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6

|||

Db 260 GYSIEE 265

RESULT 43

AD2887

dioxygenase [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)

C:Species: *Agrobacterium tumefaciens*

C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004

A:Accession: AD2887

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L. erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I ster, E.W.

A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AD2887

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-295 <KUR>

A:Cross-references: UNIPROT:Q8UCG0; GB:AE008688; PIDN:AAL43514.1; PID:g17741023; GSPDB:G

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu2527

A:Map position: circular chromosome

C:Superfamily: catechol 1,2-dioxygenase

Query Match 92.9%; Score 26; DB 2; Length 295;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6

|||

Db 10 GYFVEE 15

RESULT 44

A97663

hypothetical protein (X72850) [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)

C:Species: *Agrobacterium tumefaciens*

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004

A:Accession: A97663

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: A97663

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-295 <KUR>

A:Cross-references: UNIPROT:Q8UCG0; GB:AE007869; PIDN:AAK88258.1; PID:g15157718; GSPDB:G

C:Genetics:

A:Gene: AGR\_C\_4592

A:Map position: circular chromosome

C:Superfamily: catechol 1,2-dioxygenase





8. A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S74349  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-326 <KAN>  
A:CROSS-references: UNIPROT:Q55132; EMBL:D64001; GB:AB001339; NID:g1001102; PIDN:BAAL026  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Start codon: GTG  
C:Superfamily: conserved hypothetical protein MJ1502

Query Match 92.9%; Score 26; DB 2; Length 326;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 146 GYRVEE 151

## RESULT 50

AE1920  
GTP-binding protein [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. strain PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AE1920  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saemoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AE1920  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-337 <KUR>  
A:CROSS-references: GB:BA000019; PIDN:BAF72869.1; PID:g17130258; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr0912  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 92.9%; Score 26; DB 2; Length 337;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 332 GYRVEE 337

## RESULT 51

G64475  
GTP-binding protein, GTP1/OBG-family - Methanococcus jannaschii

C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 02-Feb-2001  
C:Accession: G64475  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Roun, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: G64475  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-354 <BUL>  
A:CROSS-references: GB:U67581; GB:L77117; NID:g1592052; PID:g1592057; TIGR:WJ1408; PID:G64475  
C:Genetics:  
A:Map position: FOR1368872-1369936  
A:Start codon: TTG

C:Superfamily: GTP-binding protein, GTP1/OBG: translation elongation factor Tu homology  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:175-293/Domain: translation elongation factor Tu homology <ETU>  
F:181-188/Region: nucleotide-binding motif A (P-loop)  
F:290-293/Region: GTP-binding NKXD motif

Query Match 92.9%; Score 26; DB 2; Length 354;  
Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 264 GYTIEE 269

## RESULT 52

T52401

branched-chain amino acid aminotransferase-like protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 09-Jul-2004  
C:Accession: T52401  
R:Kaneko, T.; Kato, T.; Sato, S.; Nakamura, Y.; Asamizu, E.; Tabata, S.  
submitted to the EMBL Data Library, September 1999

A:Reference number: Z26062  
A:Accession: T52401  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-354 <KAN>  
A:CROSS-references: UNIPROT:Q9LE06; EMBL:AP000417; PIDN:BA02558.1  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Map position: 3  
A:Introns: 11/2; 70/3; 129/3; 173/3; 205/3; 244/3; 276/3; 314/2  
C:Superfamily: branched-chain-amino-acid transaminase BAT1

Query Match 92.9%; Score 26; DB 2; Length 354;  
Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 274 GYQVEE 279

## RESULT 53

D96537

hypothetical protein F2J10.4 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: D96537  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Drenth, A.; Franks, T.; Hu, Z.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-I.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: D96537  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-356 <STO>  
A:CROSS-references: UNIPROT:Q9LPM9; GB:AE005173; NID:g8569092; PIDN:AAF76437.1; GSPDB:GN1

Qy 1 GYXVEE 6  
|||  
Db 274 GYQVEE 279

Query Match 92.9%; Score 26; DB 2; Length 356;  
Best Local Similarity 83.3%; Pred. No. 1.3e+02;

```
Matches      5;  Conservative      0;  Mismatches      1;  Indels      0;  Gaps      0;

QY      1  GYXVEE 6
      |||||
Db      275  GYQVEE 280

RESULT 54
B47411
ADPribosylarginine hydrolase (EC 3.2.2.19) - human
C:Species: Hmo sapiens (man)
C:Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 16-Aug-2004
C:Accession: B47411
R:Takada, T.; Iida, K.; Moss, J.
J. Biol. Chem. 268, 17837-17843, 1993
A:Title: Cloning and site-directed mutagenesis of human ADP-ribosylarginine hydrolase.
A:Reference number: A47411; MUID:93352593; PMID:8349667
A:Accession: B47411
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-357 <TAK>
A:Cross-references: UNIPROT:P54922; GB:L113291; NID:g402477; PIDN:AAA35555.1; PID:g402478
C:Superfamily: ADP-ribosylarginine hydrolase
C:Keywords: glycosidase; hydrolase

Query Match      92.9%;  Score 26;  DB 2;  Length 357;
Best Local Similarity 83.3%;  Pred. No. 1.3e+02;
Matches      5;  Conservative      0;  Mismatches      1;  Indels      0;  Gaps      0;

QY      1  GYXVEE 6
      |||||
Db      210  GYFVEE 215

RESULT 55
D89823
Hypothetical protein ilvE [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D89823
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yurawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:213111952; PMID:11418146
A:Accession: D89823
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <KUR>
A:Cross-references: UNIPROT:Q99W55; GB:BA0000018; PID:gl3700446; PIDN:BAB41743.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: ilvE
C:Superfamily: branched-chain-amino-acid transaminase BAT1

Query Match      92.9%;  Score 26;  DB 2;  Length 358;
Best Local Similarity 83.3%;  Pred. No. 1.3e+02;
Matches      5;  Conservative      0;  Mismatches      1;  Indels      0;  Gaps      0;

QY      1  GYXVEE 6
      |||||
Db      272  GYEVEE 277

RESULT 56
AF3184
beta-lactamase [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid AT
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AF3184
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
, erage, G.; Gilliet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
```

```
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AF3184
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <KUR>
A:Cross-references: UNIPROT:Q8UKC0; GB:AE008687; PIDN:AAL45892.1; PID:gl7743637; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu5203
A:Genome: plasmid

Query Match      92.9%;  Score 26;  DB 2;  Length 359;
Best Local Similarity 83.3%;  Pred. No. 1.3e+02;
Matches      5;  Conservative      0;  Mismatches      1;  Indels      0;  Gaps      0;

QY      1  GYXVEE 6
      |||||
Db      52  GYHVEE 57

RESULT 57
T12613
adenosylmethionine decarboxylase (EC 4.1.1.50) - common sunflower
N:Alternate names: S-adenosylmethionine decarboxylase
C:Species: Helianthus annuus (common sunflower)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C:Accession: T12613
R:Eliasson, A.; Hamann, P.; Steinmetz, A.
Submitted to the EMBL Data Library, May 1998
A:Description: Coding sequence for an S-adenosylmethionine decarboxylase from sunflower i
A:Reference number: Z17552
A:Accession: T12613
A:Status: preliminary; translated from GB/EMBL/DDBU
A:Molecule type: mRNA
A:Residues: 1-361 <ELI>
A:Cross-references: UNIPROT:O65354; EMBL:AF066078; NID:g3153905; PID:g3153906
A:Experimental source: cultivar HA300; mature pollen
C:Genetics:
A:Gene: SAD
C:Superfamily: S-adenosylmethionine decarboxylase, eukaryotic type
C:Keywords: blocked amino end; carbon-carbon lyase; carboxy-lyase
F:68/Modified site: pyruvic acid (Ser) (in mature form) #status predicted

Query Match      92.9%;  Score 26;  DB 2;  Length 361;
Best Local Similarity 83.3%;  Pred. No. 1.3e+02;
Matches      5;  Conservative      0;  Mismatches      1;  Indels      0;  Gaps      0;

QY      1  GYXVEE 6
      |||||
Db      312  GYNVEE 317

RESULT 58
E84542
hypothetical protein A2gl6650 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84542
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84542
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-369 <STO>
```

```

A;Cross-references: UNIPROT:Q9SLF2; GB:AE002093; NID:g4581132; PIDN:AAD24622.1; GSPDB:GN
C;Genetics:
A;Gene: At2g16650
A;Map position: 2

Query Match          92.9%; Score 26; DB 2; Length 369;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      165 GYVEE 170

RESULT 59
T00625
branched-chain amino acid aminotransferase homolog T2711.8 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00625
R;Fedorapfel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Araujo,
; Vyotskaia, V.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.
submitted to the EMBL Data Library, September 1998
A;Reference number: Z14193
A;Accession: T00625
A;Status: translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-384 <FED>
A;Cross-references: UNIPROT:Q93V32; EMBL:AC004122; NID:g3176693; PID:g3540185; GSPDB:GNC
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: ATSP:T2711.8
A;Map position: 1
A;Introns: 19/3; 44/2; 102/3; 161/3; 205/3; 238/3; 277/3; 309/3; 347/2
C;Superfamily: branched-chain-amino-acid transaminase BAT1

Query Match          92.9%; Score 26; DB 2; Length 384;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      307 GYKVEE 312

RESULT 60
T43170
Probable triacylglycerol lipase (EC 3.1.1.3) - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T43170
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A;Reference number: Z17323; MUID:98162722; PMID:9501991
A;Accession: T43170
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-413 <YOS>
A;Cross-references: UNIPROT:P78898; EMBL:D89249; NID:g1749705; PIDN:BAAL3910.1; PID:g174
A;Experimental source: strain PK745
C;Superfamily: triacylglycerol lipase, lingual
C;Keywords: carboxylic ester hydrolase

Query Match          92.9%; Score 26; DB 2; Length 413;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      94 GYRVEE 99

```

```

RESULT 61
G84462
probable SCARECROW gene regulator [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84462
R;Lip, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
cuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84462
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <STO>
A;Cross-references: UNIPROT:Q9S7H5; GB:AE002093; NID:g4585920; PIDN:AAD25580.1; GSPDB:GNC
C;Genetics:
A;Gene: At2g04890
A;Map position: 2

Query Match          92.9%; Score 26; DB 2; Length 413;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      387 GYAIEE 392

RESULT 62
S46801
hypothetical protein YKL027w homolog - yeast (Saccharomyces cerevisiae)
A;Alternate names: hypothetical protein YHR003C
C;Species: Saccharomyces cerevisiae
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: S46801
R;Favell, T.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of S. cerevisiae cosmid 9780.
A;Reference number: S46797
A;Accession: S46801
A;Molecule type: DNA
A;Residues: 1-429 <FAV>
A;Cross-references: UNIPROT:P38756; EMBL:U10555; NID:g500813; PID:g500821; GSPDB:GN00008;
C;Genetics:
A;Gene: MIPS:YHR003C
A;Cross-references: SGD:S0001045
A;Map position: 8R

Query Match          92.9%; Score 26; DB 2; Length 429;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      348 GYIVEE 353

RESULT 63
AE2090
site-specific DNA-methyltransferase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 16-Aug-2004
C;Accession: AE2090
R;Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Kurazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anab
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2090
A;Status: preliminary

```

A;Molecule type: DNA  
A;Residues: 1-431 <KUR>  
A;Cross-references: UNIPROT:Q8YUO9; GB:BA000019; PIDN:BAB73975.1; PID:gl7131367; GSPDB:G  
A;Experimental source: strain FCC 7120  
C;Genetics:  
C;Superfamily: Modification methylase (cytosine-specific)

Query Match 92.9%; Score 26; DB 2; Length 431;  
Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 158 GYXVEE 163

RESULT 64  
T39540  
triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T39540  
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.  
submitted to the EMBL Data Library, February 1998  
A;Reference number: Z21862  
A;Accession: T39540  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-443 <WOO>  
A;Cross-references: UNIPROT:P78898; EMBL:AL021748; PIDN:CAA16863.1; GSPDB:GN00067; SPDB:  
A;Experimental source: strain 972h; cosmid c16A3  
C;Genetics:  
A;Gene: SPDB:SPBC16A3.12c  
A;Map position: 2  
A;Introns: 11/2; 152/3; 282/1; 398/3  
C;Superfamily: triacylglycerol lipase, lingual

Query Match 92.9%; Score 26; DB 2; Length 443;  
Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 83 GYXVEE 88

RESULT 65  
E71133  
hypothetical protein PH0833 - Pyrococcus horikoshii  
C;Species: Pyrococcus horikoshii  
C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004  
C;Accession: E71133  
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A;Reference number: A71000; MUID:98344137; PMID:9679194  
A;Accession: E71133  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-483 <KAW>  
A;Cross-references: UNIPROT:O58563; GB:AP000003; NID:g3236130; PIDN:BAA29927.1; PID:g325  
A;Experimental source: strain OR3  
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
A;Gene: PH0833

Query Match 92.9%; Score 26; DB 2; Length 483;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

Db 197 GYXVEE 202  
|||  
|||

RESULT 66  
G96688  
hypothetical protein T27F4.10 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: G96688  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.;  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: G96688  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-511 <STO>  
A;Cross-references: UNIPROT:Q9C8Y3; GB:AE005173; NID:gl0092507; PIDN:AAG12907.1; GSPDB:G  
C;Genetics:  
A;Gene: T27F4.10  
A;Map position: 1

Query Match 92.9%; Score 26; DB 2; Length 511;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 480 GYXVEE 485

RESULT 67  
T51475  
RGA-like protein - Arabidopsis thaliana  
N;Alternate names: protein K3M16\_60  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: T51475  
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew  
submitted to the Protein Sequence Database, August 2000  
A;Reference number: Z25394  
A;Accession: T51475  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-523 <SAT>  
A;Cross-references: UNIPROT:Q9LF53; EMBL:AL391150  
A;Experimental source: cultivar Columbia; BAC clone K3M16  
C;Genetics:  
A;Map position: 5  
A;Note: K3M16\_60

Query Match 92.9%; Score 26; DB 2; Length 523;  
Best Local Similarity 83.3%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 490 GYXVEE 495

RESULT 68  
H96282  
protein F10B6.34 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: H86282  
 F;Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: H86282  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-533 <STO>  
 A;Cross-references: UNIPROT:Q9LQT8; GB:AE005172; NID:g8778219; PIDN:AAF79228.1; GSPDB:GN  
 C;Genetics:  
 A;Gene: F10B6.34  
 A;Map position: 1

Query Match 92.9%; Score 26; DB 2; Length 533;  
 Best Local Similarity 83.3%; Pred. No. 2e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 503 GYRVEE 508

## RESULT 69

S57648  
 nicotinic acetylcholine receptor - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
 C;Accession: S57648  
 R;Squire, M.D.; Tornoe, C.A.; Baylis, H.A.; Fleming, J.T.; Barnard, E.A.; Sattelle, D.B. submitted to the EMBL Data Library, April 1995  
 A;Description: Molecular cloning and functional expression of a *Caenorhabditis elegans* N A;Reference number: S57648  
 A;Accession: S57648  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-575 <SQ>  
 A;Cross-references: UNIPROT:P48182; EMBL:X86403; NID:g886857; PIDN:CAA60157.1; PID:g8868  
 C;Superfamily: acetylcholine receptor  
 C;Keywords: neurotransmitter receptor

Query Match 92.9%; Score 26; DB 2; Length 575;  
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 390 GYFVEE 395

## RESULT 70

T12896  
 probable single-strand DNA-specific exonuclease yorK - *Bacillus subtilis* phage SPBc2  
 C;Species: *Bacillus subtilis* phage SPBc2  
 C;Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
 C;Accession: T12896; D69923  
 R;Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Maue, C.; Karamata, D. submitted to the EMBL Data Library, August 1997  
 A;Description: The complete nucleotide sequence of the *Bacillus subtilis* SPBc2 prophage  
 A;Reference number: Z17583  
 A;Accession: T12896  
 A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA  
 A;Residues: 1-576 <LAZ>  
 A;Cross-references: UNIPROT:O64145; EMBL:AF020713; NID:g3025478; PID:g3025610; PIDN:AACT  
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emmerson, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997  
 A;Authors: Harwood, C.R.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, K.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Banchin, A. A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377  
 A;Accession: D69923  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-576 <KUN>  
 A;Cross-references: GB:Z99114; GB:AL009126; NID:g2634230; PIDN:CAB13927.1; PID:g2634428  
 A;Experimental source: strain 168  
 C;Genetics:  
 A;Gene: yorK  
 C;Superfamily: single-stranded-DNA-specific exonuclease RecJ

Query Match 92.9%; Score 26; DB 2; Length 576;  
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 478 GYKVEE 483

## RESULT 71

D84426  
 hypothetical protein At2g01570 [imported] - *Arabidopsis thaliana*  
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C;Accession: D84426  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A;Reference number: A84420; MUID:20083487; PMID:10617197  
 A;Accession: D84426  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-587 <STO>  
 A;Cross-references: UNIPROT:Q9SLH3; GB:AE002093; NID:g3785986; PIDN:AAAC67333.1; GSPDB:GN  
 C;Genetics:  
 A;Gene: At2g01570  
 A;Map position: 2

Query Match 92.9%; Score 26; DB 2; Length 587;  
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 555 GYRVEE 560

## RESULT 72

T27072  
 hypothetical protein Y51A2D.4 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C;Accession: T27072  
 R;McMurray, A. submitted to the EMBL Data Library, January 1998  
 A;Reference number: Z20307

```
A;Accession: T27072
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-606 <WIL>
A;Cross-references: UNIPROT:O9XXR3; EMBL:AL021497; PIDN:CAA16400.1; GSPDB:GN00023; CESP:
A;Experimental source: clone Y51A2D
C;Genetics:
A;Gene: CESP:Y51A2D.4
A;Map position: 5
A;Introns: 4/3; 76/1; 144/1; 226/3; 400/2; 446/1; 562/3
C;Superfamily: glucose transport protein

Query Match      92.9%; Score 26; DB 2; Length 606;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      |||
Db      557 GYSIEE 562

RESULT 73
F70325
conserved hypothetical protein aq_278 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: F70325
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: F70325
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-611 <AQF>
A;Cross-references: UNIPROT:O66634; GB:AE000682; NID:G2982979; PIDN:AAC06603.1; PID:G298
A;Experimental source: strain VF5
C;Genetics:
A;Gene: aq_278

Query Match      92.9%; Score 26; DB 2; Length 611;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      |||
Db      480 GYRVEE 485

RESULT 74
E69284
signal-transducing histidine kinase homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: E69284
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: E69284
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-629 <KLE>
A;Cross-references: UNIPROT:O29962; GB:AE001086; GB:AE000782; NID:G2689409; PIDN:AAB9095

Query Match      92.9%; Score 26; DB 2; Length 629;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 GYXVEE 6
      |||
Db      204 GYSIEE 209

RESULT 75
I40715
malate synthase (EC 4.1.3.2) - Corynebacterium glutamicum
C;Species: Corynebacterium glutamicum
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I40715; I40836
R;Reinscheid, D.J.; Eikmanns, B.J.; Sahm, H.
Microbiology 140, 3099-3108, 1994
A;Title: Malate synthase from Corynebacterium glutamicum: sequence analysis of the gene
A;Reference number: I40715; MUID:95111631; PMID:7812449
A;Accession: I40715
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-739 <RES>
A;Cross-references: UNIPROT:P42450; EMBL:X78491; NID:G530011; PIDN:CAA55243.1; PID:G53001
R;Lee, H.
J. Microbiol. Biotechnol. 4, 256-263, 1994
A;Title: Molecular characterization of aceB, a gene encoding malate synthase in Corynebac
A;Reference number: I40836
A;Accession: I40836
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-739 <RE2>
A;Cross-references: GB:L27123; NID:G853573; PIDN:AAA68074.1; PID:G853574
C;Genetics:
A;Gene: aceB
C;Keywords: carbon-carbon lyase; coenzyme A; oxo-acid-lyase

Query Match      92.9%; Score 26; DB 2; Length 739;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      |||
Db      106 GYLVEE 111

Search completed: November 1, 2004, 21:35:00
Job time : 49 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:15:12 ; Search time 127.5 Seconds  
(without alignments)  
27.076 Million cell updates/sec

Title: US-10-030-194A-5

Perfect score: 28

Sequence: 1 GYXVEE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	96.4	78	Q6D6P6	Q6d6p6 erwinia car
2	27	96.4	94	Q982Z2	Q982z2 rhizobium l
3	27	96.4	122	Q9AAK1	Q9aak1 caulobacter
4	27	96.4	167	Q6E432	Q6e432 nootka lupi
5	27	96.4	168	Q897V6	Q897v6 clostridium
6	27	96.4	170	Q8Z1H3	Q8z1h3 pyrobaculum
7	27	96.4	191	Q7ZUZ0	Q7zuz0 brachydanio
8	27	96.4	191	AAH71487	Aah71487 brachydan
9	27	96.4	200	Q52048	Q52048 plasmid phv
10	27	96.4	231	Q7N5F1	Q7n5f1 photorhabdu
11	27	96.4	232	Q7XYK9	Q7xyk9 chlorarachn
12	27	96.4	233	Q90Z28	Q90z28 pyrococcus
13	27	96.4	233	Q9KJX2	Q9kjsx2 myxococcus
14	27	96.4	280	1 VG85_BPPH2	P20344 bacterioph
15	27	96.4	280	1 VG85_BPPZA	P54694 staphylococ
16	27	96.4	282	1 DAAA_STAHA	P54694 staphylococ
17	27	96.4	299	Q9H1E6	Q9h1e6 thermoplas
18	27	96.4	306	2 Q9L973	Q9l973 lactococcus
19	27	96.4	324	1 GDPA_AQUAE	Q67555 aquifex aeo
20	27	96.4	326	2 Q8FAC8	Q6fac8 acinetobact
21	27	96.4	327	2 Q8PSH1	Q8psh1 methanosarc
22	27	96.4	330	2 Q8TMU0	Q8tmu0 methanosarc
23	27	96.4	343	2 Q98HP4	Q98hfa rhizobium l
24	27	96.4	344	1 SYFA_BACSU	P17921 bacillus su
25	27	96.4	344	2 Q916H0	Q916h0 pseudomonas
26	27	96.4	345	1 GBA4_DICDI	P34042 dictyosteli
27	27	96.4	347	2 Q88C43	Q88c43 pseudomonas
28	27	96.4	349	1 GBT_XENLA	P38407 xenopus lae
29	27	96.4	350	2 Q6MX97	Q6mx97 serratia ma
30	27	96.4	350	2 CAE51827	Caes1827 serratia
31	27	96.4	354	2 Q6MPJ7	Q6mpj7 bdellovibri

32	27	96.4	354	2 CAE78801	CaE78801 bdellovib
33	27	96.4	356	1 CARA_THETN	Q8bkl1 thermoanaer
34	27	96.4	356	2 Q9UVX8	Q9uvx8 emericella
35	27	96.4	356	2 Q8TFX6	Q8tfx6 aspergillus
36	27	96.4	356	2 CAF32008	Caf32008 aspergill
37	27	96.4	358	2 Q729K5	Q729k5 penicillium
38	27	96.4	363	2 Q9YD80	Q9ydx0 aeropyrum p
39	27	96.4	365	2 Q8HES7	Q8hes7 bacillus th
40	27	96.4	365	2 Q73212	Q73212 bacillus ce
41	27	96.4	365	2 Q81WF1	Q81wf1 bacillus an
42	27	96.4	365	2 AAS42835	Aas42835 bacillus
43	27	96.4	365	2 AAT33143	Aat33143 bacillus
44	27	96.4	370	1 SERC_METAC	Q8n1l1 methanosarc
45	27	96.4	373	2 Q819S2	Q819s2 bacillus ce
46	27	96.4	376	1 C1SY_PYRFU	Q53554 pyrococcus
47	27	96.4	385	2 Q8T2G2	Q8t2g2 methanopyru
48	27	96.4	385	2 Q94CF8	Q94cf8 capsicum an
49	27	96.4	390	2 Q20910	Q20910 caenorhabdi
50	27	96.4	390	2 Q72TP8	Q72tp8 leptospira
51	27	96.4	390	2 Q8F1J1	Q8f1j1 leptospira
52	27	96.4	390	2 AAS69580	Aas69580 leptospir
53	27	96.4	398	2 Q98PF1	Q98pf1 mycoplasma
54	27	96.4	413	1 NEUA_STRA3	Q53598 streptococc
55	27	96.4	413	1 NEUA_STRA5	Q9afg9 streptococc
56	27	96.4	413	2 Q9ALW4	Q9alw4 streptococc
57	27	96.4	413	2 Q93T10	Q93t10 streptococc
58	27	96.4	431	2 Q7MLQ2	Q7mlq2 bordetella
59	27	96.4	447	2 Q8AB03	Q8ab03 bacteroides
60	27	96.4	457	1 CGPB_FUSSO	Q00858 fusarium so
61	27	96.4	466	2 Q9CH26	Q9ch26 lactococcus
62	27	96.4	466	2 Q9CH26	Q9ch26 lactococcus
63	27	96.4	528	2 Q8SVJ8	Q8svj8 encephalito
64	27	96.4	553	1 HUTU_STAAM	Q93lg1 staphylococ
65	27	96.4	553	1 HUTU_STAAN	Q99r2 staphylococ
66	27	96.4	553	2 Q6G6Y9	Q6g6y9 staphylococ
67	27	96.4	575	2 Q84Q92	Q84q92 oryza sativ
68	27	96.4	645	2 Q7S7Z0	Q7s7z0 neurospora
69	27	96.4	646	2 Q8YT29	Q8yt29 anabaena sp
70	27	96.4	653	2 Q9C2K2	Q9c2k2 neurospora
71	27	96.4	751	2 Q88EF1	Q88ef1 pseudomonas
72	27	96.4	861	1 GCR3_YEAST	P34160 saccharomyc
73	27	96.4	877	2 Q8KQV6	Q8kqv6 vibrio chol
74	27	96.4	879	2 Q8ZS90	Q8zs90 anabaena sp
75	27	96.4	885	2 Q84A02	Q84a02 escherichia
76	27	96.4	894	2 Q8RL18	Q8rl18 providencia
77	27	96.4	1035	2 Q6NBZ7	Q6nbz7 rhodopseude
78	27	96.4	1035	2 CAE26122	CaE26122 rhodopseu
79	27	96.4	1050	2 Q8PU41	Q8pu41 methanosarc
80	27	96.4	1075	2 Q9H2G1	Q9h2g1 homo sapien
81	27	96.4	1139	2 Q8NF92	Q8nf92 homo sapien
82	27	96.4	1165	2 Q72TU4	Q72tu4 leptospira
83	27	96.4	1165	2 Q8F1C9	Q8f1c9 leptospira
84	27	96.4	1165	2 AAS69534	Aas69534 leptospir
85	27	96.4	1229	2 Q82Y40	Q82y40 nitrosomona
86	27	96.4	1374	2 Q82148	Q82148 arabidopsis
87	27	96.4	1868	2 Q6H242	Q6h242 lactobacill
88	27	96.4	1978	2 Q14966	Q14966 homo sapien
89	27	96.4	1978	2 Q723T7	Q723t7 homo sapien
90	27	96.4	2241	2 Q7XRJ2	Q7xrj2 oryza sativ
91	27	96.4	3597	2 Q81LR5	Q81lr5 plasmodium
92	26	92.9	64	2 Q8TVE0	Q8tve0 methanopyru
93	26	92.9	77	2 Q96WB9	Q96wb9 fusarium sp
94	26	92.9	79	2 Q8U493	Q8u493 pyrococcus
95	26	92.9	80	2 Q9UXW4	Q9uxw4 pyrococcus
96	26	92.9	95	2 Q25553	Q25553 helicobacte
97	26	92.9	95	2 Q9ZKV3	Q9zkv3 helicobacte
98	26	92.9	102	2 Q74MG9	Q74mg9 nanoarchaea
99	26	92.9	102	2 AAR38881	Aar38881 nanoarcha
100	26	92.9	116	1 Y053_ENTFA	Q831p7 enterococcu
101	26	92.9	116	2 Q833J9	Q833j9 enterococcu
102	26	92.9	116	2 Q92FF9	Q92ff9 listeria in
103	26	92.9	116	2 Q8YAL8	Q8yal8 listeria mo
104	26	92.9	116	2 Q724V6	Q724v6 listeria mo

105	26	92.9	116	2	AAT02905	Aat02905 listeria	178	319	1	YHAI_CRYPA	P10941 cryphonectr
106	26	92.9	124	2	Q87B26	Q87b26 xylella fas	179	324	1	ERA_ANASP	Q8yy08 anabaena sp
107	26	92.9	124	2	Q9PG56	Q9pg56 xylella fas	180	325	1	RPOA_MYCGA	Q9rdv6 mycoplasma
108	26	92.9	125	2	Q7PFI1	Q7pfi1 anopheles g	181	326	1	RNZ_SYNY3	Q55132 synechocyst
109	26	92.9	132	2	Q8XK66	Q8kke6 clostridium	182	326	1	Q70IX7	Q70ix7 streptomyc
110	26	92.9	137	1	FRXA_PYRFU	Q51795 pyrococcus	183	339	2	CAE22477	Cae22477 streptomy
111	26	92.9	137	2	Q7X4Y1	Q7x4y1 rhodospiril	184	343	2	Q6L743	Q6l743 streptomyc
112	26	92.9	139	2	Q8K322	Q8k3z2 mus musculu	185	343	2	BAD20754	Bad20754 streptomy
113	26	92.9	143	2	Q8S363	Q8s363 anisocarpus	186	350	1	YE08_METJA	Q58803 methanococ
114	26	92.9	146	2	Q88Y19	Q88yt3 lactobacill	187	354	1	BCA4_ARATH	Q9le06 arabidopsis
115	26	92.9	151	2	Q58407	Q88407 pyrococcus	188	354	1	LEU3_BACCR	Q81g11 bacillus ce
116	26	92.9	153	2	Q8PFG8	Q8pgp8 xanthomonas	189	356	1	BCAC_HUMAN	Q9lpn9 arabidopsis
117	26	92.9	154	2	Q9XJU4	Q9xju4 streptococ	190	357	1	ARH1_HUMAN	P54922 homo sapien
118	26	92.9	155	2	Q8PC55	Q8pcs5 xanthomonas	191	357	2	Q74FF3	Q74ff3 geobacter s
119	26	92.9	161	2	Q8Y9R1	Q8y9r1 listeria mo	192	357	2	AAR33986	Aar33986 geobacter
120	26	92.9	165	2	Q8XJDS	Q8xjd5 streptococ	193	358	1	ILVE_STAAM	Aar33986 geobacter
121	26	92.9	165	2	Q94M58	Q94m58 streptococ	194	358	1	ILVE_STAEP	Q89w55 staphyloc
122	26	92.9	166	2	Q9MCHO	Q9mch0 streptococ	195	358	2	Q6G5T3	Q8cg78 staphyloc
123	26	92.9	167	2	Q8MCH9	Q8mch9 streptococ	196	358	2	Q6GJJB4	Q8gb34 staphyloc
124	26	92.9	171	2	Q8MXG6	Q8mxg6 caenorhabdi	197	359	2	Q7D3N4	Q8gjb4 staphyloc
125	26	92.9	177	2	Q22434	Q22434 nicotiana t	198	359	2	Q9F329	Q7d3n4 agrobacteri
126	26	92.9	180	2	Q8X4M0	Q8x4m0 lactobacill	199	359	2	Q8UKC0	Q9f329 staphyloc
127	26	92.9	181	2	Q34051	Q34051 streptococ	200	361	1	DCAM_DAUCA	Q8ukc0 agrobacteri
128	26	92.9	182	2	Q61QZ6	Q61qz6 mus musculu	201	361	1	DCAM_HELAN	Q9axe3 daucus caro
129	26	92.9	182	2	AAH71245	Aah71245 mus muscu	202	361	2	Q6RUQ3	Q65354 helianthus
130	26	92.9	195	2	Q9VD69	Q9v6d9 drosophila	203	361	2	AAR84406	Q6r3q3 daucus caro
131	26	92.9	201	2	Q6ZT50	Q6zt50 homo sapien	204	361	2	AAR84408	Aar84408 daucus ca
132	26	92.9	201	2	BAC86744	Bac86744 homo sapi	205	362	2	Q8CUM8	Q8cun8 oceanobacil
133	26	92.9	204	1	REX2_THEMEA	Q9xiel thermotoga	206	362	2	Q7PTH8	Q8cun8 oceanobacil
134	26	92.9	207	2	Q7M2D4	Q7m2d4 triticum ti	207	362	2	Q9H9F3	Q7pth8 anopheles g
135	26	92.9	210	2	Q3FJL8	Q3fjl8 arabidopsis	208	362	2	Q9H9F3	Q9h9f3 homo sapien
136	26	92.9	223	2	Q801A6	Q801a6 latimeria m	209	366	2	Q9SLF2	Q9slf2 arabidopsis
137	26	92.9	224	1	REX_THETN	Q8rc93 thermoanaer	210	369	2	Q9SLF2	P52878 methanosarc
138	26	92.9	231	2	Q8U2A9	Q8u2a9 pyrococcus	211	371	2	Q8PV68	Q8pv68 methanosarc
139	26	92.9	235	2	Q7QH00	Q7qhu0 anopheles g	212	371	2	Q79VE5	Q79ve5 corynebacte
140	26	92.9	235	2	Q6WRX8	Q6wrx8 enterococcu	213	379	2	CAF20544	Caf20544 corynebac
141	26	92.9	235	2	Q9KHM3	Q9khm3 enterococcu	214	381	2	Q7V3C0	Q7v3c0 prochloroco
142	26	92.9	235	2	AAQ16268	AAq16268 enterococ	215	382	2	Q9FCP4	Q9fcpc rhizobium l
143	26	92.9	242	2	Q92DH2	Q92dh2 listeria in	216	383	2	BCAL_ARATH	Q9fcpc rhizobium l
144	26	92.9	242	2	Q8Y8P8	Q8y8p8 listeria mo	217	384	1	Q8DL05	Q8dl05 synechococ
145	26	92.9	242	2	Q721W9	Q72iw9 listeria mo	218	386	1	CTBP_DROME	O46036 drosophila
146	26	92.9	242	2	AAQ3645	Aat03645 listeria	219	386	2	Q7SPF2	Q7sfr2 neurospora
147	26	92.9	244	1	T2E5_ECOLI	P04390 escherichia	220	386	2	Q93PA6	Q93pa6 microscilla
148	26	92.9	245	2	AAQ18131	AAq18131 escherich	221	386	2	AAQ39315	AAq39315 drosophil
149	26	92.9	252	2	Q9KCS0	Q9kcs0 bacillus ha	222	387	2	Q99JU4	Q99ju4 mus musculu
150	26	92.9	252	2	V8PB_SOYEN	P10743 glycine max	223	388	2	Q732P8	Q732p8 mycobacteri
151	26	92.9	257	1	Q80854	O80854 arabidopsis	224	388	2	AAQ03870	AAq03870 mycobacte
152	26	92.9	257	2	Q83EE5	Q83ees coxiella bu	225	397	2	Q70KX9	Q70kx9 hordeum vul
153	26	92.9	257	2	AAQ62413	AAq62413 arabidops	226	397	2	CAE00460	Cae00460 hordeum v
154	26	92.9	270	2	Q73DA9	Q73da9 bacillus ce	227	398	1	ACK1_CLOPE	Q8xmw5 clostridium
155	26	92.9	270	2	Q6R7C2	Q6r7c2 ostreid her	228	398	2	Q99K70	Q99k70 mus musculu
156	26	92.9	270	2	AAQ00993	AAq00993 ostreid h	229	399	2	Q9HB90	Q9hb90 homo sapien
157	26	92.9	270	2	AAQ39736	AAq39736 bacillus	230	399	2	Q9H202	Q9h202 homo sapien
158	26	92.9	275	1	YK42_SYNY3	P73079 synechocyst	231	399	2	Q8DZN9	Q8dzn9 streptococ
159	26	92.9	278	2	Q9SR59	Q9srs9 arabidopsis	232	399	2	Q8E5D7	Q8e5d7 streptococ
160	26	92.9	285	1	YFTE_BACSU	P54721 bacillus su	233	411	2	Q8FNF9	Q8fnp9 corynebacte
161	26	92.9	287	2	Q6PFM6	Q6pfm6 brachydanio	234	413	2	Q9S7H5	Q9s7h5 arabidopsis
162	26	92.9	287	2	AAH57492	AAh57492 brachydan	235	415	2	Q8GRH8	Q8grh8 oryza sativ
163	26	92.9	288	2	Q8CUV6	Q8cuv6 oceanobacil	236	421	2	Q93LL9	Q93ll9 nostoc punc
164	26	92.9	291	2	Q6DDU1	Q6ddul xenopus lae	237	429	1	YHG3_YEAST	Yhg3 yeast
165	26	92.9	295	2	Q7CWU7	Q7cwu7 agrobacteri	238	431	2	Q8YUQ9	Q8yqu9 anabaena sp
166	26	92.9	295	2	Q6MMJ8	Q6mmj8 bdellovibri	239	433	2	Q7QBX7	Q7qpx7 anopheles g
167	26	92.9	295	2	Q8UCG0	Q8ucg0 agrobacteri	240	433	2	CAF29987	Caf29987 methanoco
168	26	92.9	295	2	CAE79506	Cae79506 bdellovib	241	433	2	Q6FN42	Q6fn42 candida gla
169	26	92.9	298	2	Q9L7R2	Q9l7r2 salmonella	242	438	2	P78898	P78898 schizosacch
170	26	92.9	298	2	Q7CFC9	Q7cpc9 salmonella	243	443	2	Q8TWS5	Q8tws5 methanopyru
171	26	92.9	299	2	Q9V6F0	Q9v6f0 drosophila	244	445	2	Q8TWS5	Q8tws5 methanopyru
172	26	92.9	303	2	Q82G06	Q82g06 streptomyce	245	446	2	Q7WUI7	Q7wui7 thermus the
173	26	92.9	304	2	Q6NRU5	Q6nru5 xenopus lae	246	465	2	Q7V108	Q7v108 prochloroco
174	26	92.9	304	2	AAH70619	AAh70619 xenopus l	247	468	2	Q8WQM4	Q8wqm4 plasmodium
175	26	92.9	311	1	RNZ_LACPL	Q88vg6 lactobacill	248	472	2	O15501	O15501 homo sapien
176	26	92.9	313	2	Q23549	Q23549 arabidopsis	249	472	2	Q8TMS5	Q8tms5 wolbachia s
177	26	92.9	315	1	ERA_SYNY3	Q55526 synechocyst	250	475	2	Q75QM9	Q75qm9 bacterioph



251	26	92.9	475	2	BAD16786	Bad16786 bacteriop	324	92.9	610	2	O6FC42	O6fc42 acinetobact
252	26	92.9	479	2	O6C125	O6c125 yarrowia li	325	92.9	611	2	O66634	O66634 aquifex aeo
253	26	92.9	480	2	O8K1E9	O8k1e9 mus musculu	326	92.9	618	2	O8W127	O8w127 hordeum vul
254	26	92.9	483	2	O58563	O58563 pyrococcus	327	92.9	623	2	O9ST59	O9st59 triticum ae
255	26	92.9	489	2	O8PT52	O8pt52 methanosarc	328	92.9	625	2	O7G7J6	O7g7j6 oryza sativ
256	26	92.9	489	2	O8PT53	O8pt53 methanosarc	329	92.9	625	2	O9MB96	O9mb96 oryza sativ
257	26	92.9	497	2	P90338	P90338 saguaro cac	330	92.9	629	2	O29962	O29962 archaeoglob
258	26	92.9	502	1	C7ZG_ARATH	C7lzm7 arabidopsis	331	92.9	630	2	Q9ST48	Q9st48 zea mays (m
259	26	92.9	503	2	O739A0	O739a0 bacillus ce	332	92.9	633	2	Q8BMY1	Q8bmy1 m mus muscu
260	26	92.9	503	2	O81E28	O81e28 bacillus ce	333	92.9	650	2	Q617K8	Q617k8 myxococcus
261	26	92.9	503	2	AAe41162	AAe41162 bacillus	334	92.9	662	2	O7Q8W8	O7q8w8 anopheles g
262	26	92.9	511	2	O95J54	O95j54 macaca fasc	335	92.9	662	2	O65367	O65367 arabidopsis
263	26	92.9	511	2	O9C8Y3	O9c8y3 arabidopsis	336	92.9	693	2	Q9JJ99	Q9jj99 m mus muscu
264	26	92.9	513	2	O8S1F1	O8s1f1 oryza sativ	337	92.9	698	2	Q8NC60	Q8nc60 homo sapien
265	26	92.9	518	2	O9L1H4	O9l1h4 streptomyce	338	92.9	698	2	Q9BSQ9	Q9bsq9 homo sapien
266	26	92.9	523	2	O9LF53	O9lf53 arabidopsis	339	92.9	707	2	O6C1X2	O6c1x2 yarrowia li
267	26	92.9	524	2	O8FRK7	O8frk7 corynebacte	340	92.9	718	2	O8GX10	O8gx10 arabidopsis
268	26	92.9	526	2	O7MX62	O7mx62 porphyronon	341	92.9	727	2	O6ZLD6	O6zld6 oryza sativ
269	26	92.9	532	1	NAD6_RHINE	Q92r32 rhizobium m	342	92.9	727	2	BAC83029	Bac83029 oryza sat
270	26	92.9	532	2	O23643	O23643 arabidopsis	343	92.9	738	1	MAS2_CORGL	P42450 corynebacte
271	26	92.9	532	2	O23724	O23724 arabidopsis	344	92.9	752	2	Q8NAN7	Q8nan7 homo sapien
272	26	92.9	533	2	O9LQT8	O9lqt8 arabidopsis	345	92.9	755	2	P89107	P89107 saguaro cac
273	26	92.9	534	2	O8S357	O8s357 dubautia ar	346	92.9	773	2	O65987	O65987 cardamine c
274	26	92.9	534	2	O8CMK3	O8cmk3 staphylococ	347	92.9	779	2	Q9XB99	Q9xbp9 myxococcus
275	26	92.9	535	2	Q8S369	Q8s369 madia sativ	348	92.9	790	2	Q02744	Q02744 pneumocysti
276	26	92.9	536	2	O8S374	O8s374 dubautia me	349	92.9	803	2	O6CWS7	Q6cws7 kluyveromyc
277	26	92.9	536	2	O8S375	O8s375 dubautia ci	350	92.9	824	2	Q6UW79	Q6uw79 oryza sativ
278	26	92.9	536	2	O8S376	O8s376 dubautia kn	351	92.9	824	2	AAQ56498	AAq56498 oryza sat
279	26	92.9	537	2	O84TQ7	O84tq7 gossypium h	352	92.9	836	2	O7CZS6	O7czs6 agrobacteri
280	26	92.9	537	2	O8S353	O8s353 dubautia mi	353	92.9	836	2	Q8UG47	Q8ug47 agrobacteri
281	26	92.9	537	2	O8S355	O8s355 dubautia me	354	92.9	852	2	Q8IKW0	Q8ikw0 plasmodium
282	26	92.9	537	2	O8S356	O8s356 dubautia kn	355	92.9	863	2	O29837	O29837 archaeoglob
283	26	92.9	537	2	O8S370	O8s370 argyroxiphi	356	92.9	873	1	LDVR_HUMAN	P98155 homo sapien
284	26	92.9	538	2	O8RUC4	O8ruc4 wilkesia gy	357	92.9	878	2	Q8EDN8	Q8edn8 shewanella
285	26	92.9	538	2	O8S371	O8s371 argyroxiphi	358	92.9	881	2	O6FET8	O6fet8 acinetobact
286	26	92.9	538	2	O8S372	O8s372 argyroxiphi	359	92.9	905	2	O67483	O67483 banzi virus
287	26	92.9	538	2	O8S373	O8s373 dubautia ra	360	92.9	909	2	Q9SIR8	Q9sik8 arabidopsis
288	26	92.9	538	2	O8S377	O8s377 dubautia ar	361	92.9	910	1	DRL8_ARATH	O8w3k3 arabidopsis
289	26	92.9	539	2	O8S354	O8s354 dubautia ra	362	92.9	910	1	VILI_ARATH	O816k3 arabidopsis
290	26	92.9	540	2	O8S365	O8s365 calycadenia	363	92.9	929	2	O8MY79	O8my79 haemaphysal
291	26	92.9	540	2	O8S378	O8s378 dubautia ar	364	92.9	934	2	O81ZJ1	O81zj1 homo sapien
292	26	92.9	541	2	O8S361	O8s361 argyroxiphi	365	92.9	934	2	O6PFH0	O6pfh0 mus musculu
293	26	92.9	542	2	O8S360	O8s360 argyroxiphi	366	92.9	934	2	AAH57560	AAh57560 mus muscu
294	26	92.9	542	2	O8S362	O8s362 argyroxiphi	367	92.9	943	2	Q8JGT4	Q8jgt4 xenopus lae
295	26	92.9	544	2	O8S359	O8s359 argyroxiphi	368	92.9	945	2	O86SN3	O86sn3 homo sapien
296	26	92.9	547	2	O8CXW1	O8cxw1 arabidopsis	369	92.9	945	2	O08722	O08722 rattus norv
297	26	92.9	547	2	O9SRP9	O9srp9 arabidopsis	370	92.9	945	2	O8K1S3	O8k1s3 mus musculu
298	26	92.9	548	2	O8S367	O8s367 carliquistia	371	92.9	945	2	Q9D398	Q9d398 mus musculu
299	26	92.9	548	2	O8S368	O8s368 carliquistia	372	92.9	945	2	AAQ88717	AAq88717 homo sapi
300	26	92.9	551	2	O8T294	O8t294 dictyosteli	373	92.9	1008	2	Q80Y85	Q80y85 mus musculu
301	26	92.9	554	2	O8N1Y2	O8n1y2 homo sapien	374	92.9	1015	2	Q82LB5	Q82lb5 streptomyce
302	26	92.9	562	2	O8RB33	O8rb33 thermoanaer	375	92.9	1021	2	O8YSB9	O8ysb9 anabaena ep
303	26	92.9	575	1	ACH4_CAEBL	P48182 caenorhabdi	376	92.9	1040	2	O7P6Q0	O7p6q0 fuobacteri
304	26	92.9	576	2	O64145	O64145 bacterioph	377	92.9	1051	2	Q01694	Q01694 pneumocysti
305	26	92.9	576	2	O31903	O31903 bacillus su	378	92.9	1058	1	CARB_STRP3	Q8k7y3 streptococc
306	26	92.9	579	2	O6E106	O6e106 cucurbita m	379	92.9	1058	1	CARB_STRP8	P58941 streptococc
307	26	92.9	587	2	O23642	O23642 arabidopsis	380	92.9	1058	1	CARB_STRPY	Q9a0c6 streptococc
308	26	92.9	587	2	O23725	O23725 arabidopsis	381	92.9	1059	2	O8DUF3	O8dup3 streptococc
309	26	92.9	587	2	O9SLH3	O9slh3 arabidopsis	382	92.9	1060	2	C834E2	O834e2 enterococcu
310	26	92.9	587	2	Q941D4	Q941d4 arabidopsis	383	92.9	1062	1	CARB_EACHD	Q9k9v9 bacillus ha
311	26	92.9	587	2	O6E105	O6e105 cucurbita m	384	92.9	1087	1	XYNX_CLOTH	P38535 clostridium
312	26	92.9	587	2	AAQ65090	AAq65090 arabidops	385	92.9	1138	1	DRL9_ARATH	Q8w3k0 arabidopsis
313	26	92.9	588	2	O7Y1B6	O7y1b6 lycopersico	386	92.9	1157	1	XYNA_THESA	P36917 thermoanaer
314	26	92.9	590	2	O8S4W7	O8s4w7 vitis vinif	387	92.9	1187	2	O93284	O93284 fugu rubrip
315	26	92.9	590	2	O8W5P5	O8w5p5 oryza sativ	388	92.9	1234	2	O60046	O60046 thermoanaer
316	26	92.9	591	2	O8AVF5	O8avf5 xenopus lae	389	92.9	1237	2	Q7SEH3	Q7seh3 neurospora
317	26	92.9	593	2	O814Z3	O814z3 bacillus ce	390	92.9	1312	2	Q8MVR0	Q8mvr0 dictyosteli
318	26	92.9	595	2	O7XDF5	O7xdf5 oryza sativ	391	92.9	1329	2	O27416	O27416 methanobact
319	26	92.9	595	2	O8W5P8	O8w5p8 oryza sativ	392	92.9	1348	2	O60043	O60043 thermoanaer
320	26	92.9	596	2	O7Q0L9	O7q0l9 anopheles g	393	92.9	1349	2	O8TT86	O8tt86 methanosarc
321	26	92.9	603	2	O7Q2X11	O7qx11 bacillus ce	394	92.9	1379	2	O07893	O07893 antirrhinum
322	26	92.9	603	2	AA844295	AA844295 bacillus	395	92.9	1380	2	Q40001	Q40001 hordeum vul
323	26	92.9	606	2	Q9XXR3	Q9xxr3 caenorhabdi	396	92.9	1381	2	Q39049	Q39049 arabidopsis

397	26	92.9	1381	2	Q9FNB0	Q9fnb0 arabidopsis	470	25	89.3	306	2	Q6F9V3	Q6f9v3 acinetobact
398	26	92.9	1381	2	Q94C01	Q94c01 hordeum vul	471	25	89.3	313	1	RNZ_ENTFA	Q834g2 enterococc
399	26	92.9	1381	2	Q8RY14	Q8ry14 arabidopsis	472	25	89.3	315	2	Q55788	Q55788 synchocyst
400	26	92.9	1382	2	Q22435	Q22435 nicotiana t	473	25	89.3	317	2	Q93445	Q93445 oryzias lat
401	26	92.9	1383	2	Q65808	Q65808 glycine max	474	25	89.3	318	2	Q73P13	Q73p13 treponema d
402	26	92.9	1395	1	LCTA_LATWA	Q9xzc0 latrodectus	475	25	89.3	318	2	Q7S219	Q7s219 xenopus lae
403	26	92.9	1653	2	QUGO1	Q9ug01 homo sapien	476	25	89.3	318	2	AAS11477	Aas11477 treponema
404	26	92.9	1743	1	TAGC_DICDI	Q23868 dictyosteli	477	25	89.3	323	1	ANX5_CYNPY	P70075 cynops pyrr
405	26	92.9	1749	2	Q6NZK1	Q6nzk1 mus musculu	478	25	89.3	323	2	Q6GR36	Q6gr36 xenopus lae
406	26	92.9	1749	2	Q6VH22	Q6vhn2 mus musculu	479	25	89.3	323	2	Q6DDN9	Q6ddn9 xenopus lae
407	26	92.9	1749	2	Q9JKU3	Q9jku3 rattus norv	480	25	89.3	324	2	Q9HVR6	Q9hvr6 pseudomonas
408	26	92.9	1749	2	AAR053390	Aar05390 mus muscu	481	25	89.3	331	2	Q92EU2	Q92eu2 listeria in
409	26	92.9	1749	2	AAH66096	Aah66096 mus muscu	482	25	89.3	331	2	Q8YA18	Q8ya18 listeria mo
410	26	92.9	1905	1	TAGB_DICDI	P54683 dictyosteli	483	25	89.3	331	2	Q723V7	Q723v7 listeria mo
411	26	92.9	2143	2	Q80T20	Q80t20 mus musculu	484	25	89.3	331	2	AAT031154	Aat031154 listeria
412	26	92.9	2345	2	Q7QTH1	Q7qth1 giardia lam	485	25	89.3	333	2	Q9UXF1	Q9uxf1 sulfolobus
413	26	92.9	2862	2	Q8R874	Q8r874 thermoanaer	486	25	89.3	336	2	Q97VM0	Q97vm0 sulfolobus
414	26	92.9	2862	2	Q983H6	Q983h6 thizobium l	487	25	89.3	339	2	Q97IF6	Q97if6 clostridium
415	25	89.3	26	2	Q52919	O52919 campylobact	488	25	89.3	351	2	Q8TZC7	Q8tzc7 methanopyru
416	25	89.3	71	2	Q6CB27	Q6cb27 varrowia li	489	25	89.3	352	2	Q6CR87	Q6crr87 kluyveromyc
417	25	89.3	82	2	Q7VLB8	Q7vlb8 haemophilus	490	25	89.3	359	2	Q6CR86	P96085 thermoplasm
418	25	89.3	140	2	Q6XKD7	Q6xkd7 sulfolobus	491	25	89.3	360	2	P96085	Q9hi56 thermoplasm
419	25	89.3	144	2	Q9X154	Q9xi54 thermotoga	492	25	89.3	360	2	Q70WQ3	Q70wq3 thermoprote
420	25	89.3	150	2	Q05969	Q05969 brassica ol	493	25	89.3	374	2	CAD56502	Cad56502 thermopro
421	25	89.3	155	2	Q7P4B3	Q7p4b3 fusobacteri	494	25	89.3	376	1	YQGT_BACSU	P54497 bacillus su
422	25	89.3	159	2	Q6MTJ1	Q6mtj1 mycoplasma	495	25	89.3	379	1	CBID_SALTY	O8z5m8 salmonella
423	25	89.3	164	2	Q8RHG7	Q8rhg7 fusobacteri	496	25	89.3	379	1	CBID_SALTY	O05628 salmonella
424	25	89.3	180	2	Q26058	O26058 helicobacte	497	25	89.3	383	2	Q92H41	Q92h41 listeria in
425	25	89.3	180	2	Q9ZJ95	Q9zj95 helicobacte	498	25	89.3	383	2	Q7QU72	Q7qu72 giardia lam
426	25	89.3	182	2	Q6NH08	Q6nh08 corynebacte	499	25	89.3	398	2	Q9V796	Q9v796 drosophila
427	25	89.3	182	2	CNE49867	Cae49867 corynebac	500	25	89.3	398	2	Q9V796	Q9v796 drosophila
428	25	89.3	183	2	Q8XNA9	O8xna9 clostridium	501	25	89.3	399	1	YF07_METJA	Q58902 methanococc
429	25	89.3	183	2	Q8XMA9	Q8xma9 drosophila	502	25	89.3	409	2	Q8ZWP2	Q8zwp2 pyrobaculum
430	25	89.3	196	2	Q6XHM0	Q6xhm0 drosophila	503	25	89.3	410	2	Q8A5Q2	Q8a5q2 bacteroides
431	25	89.3	202	1	ARI0186	Aar10186 drosophil	504	25	89.3	412	1	HISX_THET2	P62460 thermus the
432	25	89.3	223	2	HIS5_LACLA	Q02132 lactococcus	505	25	89.3	412	2	AAS80718	Aas80718 thermus t
433	25	89.3	234	2	Q7RVR6	Q7rvr6 neurospora	506	25	89.3	416	2	Q8H8S3	Q8h8s3 oryza sativ
434	25	89.3	234	2	Q74N42	Q74n42 nanoarchaeu	507	25	89.3	419	1	YDEA_SCHPO	Q10443 schizosacch
435	25	89.3	234	2	O05377	O05377 actinobacil	508	25	89.3	419	1	Q6TG03	Q6tg03 campylobact
436	25	89.3	234	2	Q9AQB2	Q9aqb2 actinobacil	509	25	89.3	419	2	Q7BPR7	Q7bpr7 campylobact
437	25	89.3	234	2	Q9AQB2	Aar38875 nanoarcha	510	25	89.3	419	2	AAR99171	Aar99171 campyloba
438	25	89.3	235	2	Q9C2Q6	Q9c2q6 neurospora	511	25	89.3	419	2	AAR82887	Aar82887 campyloba
439	25	89.3	240	2	Q81DB2	Q81db2 bacillus ce	512	25	89.3	429	2	Q7XL25	Q7xl25 oryza sativ
440	25	89.3	240	2	Q8DP95	Q8dp95 streptococc	513	25	89.3	432	2	Q9PND9	Q9pnd9 campylobact
441	25	89.3	241	2	Q7P2X6	Q7p2x6 fusobacteri	514	25	89.3	434	1	GATA_METJA	Q58560 methanococc
442	25	89.3	242	2	Q8LJD19	O8ldi9 arabidopsis	515	25	89.3	442	2	Q8XJL5	Q8xjl5 clostridium
443	25	89.3	242	2	Q97Q07	Q97q07 streptococc	516	25	89.3	469	2	Q18644	Q18644 caenorhabdi
444	25	89.3	244	2	Q91M93	Q91m93 arabidopsis	517	25	89.3	470	2	Q6SED7	Q6sed7 lactobacill
445	25	89.3	246	2	Q59242	O59242 pyrococcus	518	25	89.3	470	2	AAR27374	Aar27374 lactobaci
446	25	89.3	249	2	Q9V171	Q9v171 pyrococcus	519	25	89.3	470	2	AAS09206	Aas09206 lactobaci
447	25	89.3	249	2	Q8U0H1	Q8u0h1 pyrococcus	520	25	89.3	480	2	Q9LXS7	Q9lxs7 arabidopsis
448	25	89.3	250	2	Q6CY76	Q6cy76 kluyveromyc	521	25	89.3	485	2	Q6Z684	Q6z684 oryza sativ
449	25	89.3	255	2	O51484	O51484 borrelia bu	522	25	89.3	485	2	BAD17253	Bad17253 oryza sat
450	25	89.3	256	2	Q9LWMS	Q9lmw5 arabidopsis	523	25	89.3	491	2	Q8TEK5	Q8tek5 homo sapien
451	25	89.3	257	2	Q8RHK2	Q8rnh2 fusobacteri	524	25	89.3	496	2	Q7SYD7	Q7syd7 brachydanio
452	25	89.3	264	2	Q8PZS8	Q8pzs8 methanosarc	525	25	89.3	503	2	Q6HJF1	Q6hjf1 bacillus th
453	25	89.3	266	2	Q18643	Q18643 caenorhabdi	526	25	89.3	503	2	Q81R73	Q81r73 bacillus an
454	25	89.3	266	2	P93536	P93536 sophora jap	527	25	89.3	503	2	AAT31299	Aat31299 bacillus
455	25	89.3	267	1	Y555_AQUAE	O66830 aquifex aeo	528	25	89.3	504	2	Q94568	Q94568 galliera me
456	25	89.3	283	2	Q9AZ90	Q9az90 lactobacill	529	25	89.3	504	2	Q94568	Q94568 galliera me
457	25	89.3	286	1	RT28_YEAST	P21771 saccharomyc	530	25	89.3	509	2	Q82NM3	Q82nm3 streptomyc
458	25	89.3	286	2	AAS56069	Aas56069 saccharom	531	25	89.3	514	2	Q82NM3	Q82nm3 streptomyc
459	25	89.3	289	2	Q8VVQ6	Q8vvq6 staphylococ	532	25	89.3	514	2	Q8E481	Q8e481 streptococc
460	25	89.3	300	1	TF2B_PYRBU	Q9v0v5 pyrococcus	533	25	89.3	514	2	Q8DHW9	Q8dhw9 synchococc
461	25	89.3	300	1	TF2B_PYRBU	P61998 pyrococcus	534	25	89.3	552	2	HEMA_IAGUA	P13101 influenza a
462	25	89.3	300	1	TF2B_PYRBU	O59151 pyrococcus	535	25	89.3	566	1	HEMA_IAGU2	P13103 influenza a
463	25	89.3	302	1	TF2B_PYRBU	P61999 pyrococcus	536	25	89.3	566	1	HEMA_IAPIL	P13102 influenza a
464	25	89.3	302	2	Q6ZES1	Q6zes1 synchocyst	537	25	89.3	586	2	Q8H612	Q8h612 oryza sativ
465	25	89.3	305	1	BAD01829	Bad01829 synchocyst	538	25	89.3	602	2	Q86KU7	Q86ku7 dictyosteli
466	25	89.3	305	1	RNZ_EC057	Q8xcz0 escherichia	539	25	89.3	624	2	Q8SZ15	Q8sz15 drosophila
467	25	89.3	305	1	RNZ_EC016	Q8ffk8 escherichia	540	25	89.3	641	2	Q6TUY6	Q6tuy6 yaba monkey
468	25	89.3	305	1	RNZ_EC016	Q47012 escherichia	541	25	89.3	641	2	AAR07383	Aar07383 yaba monk
469	25	89.3	305	1	TF2B_PYRKO	P58109 pyrococcus	542	25	89.3	661	2	Q86YD2	Q86yd2 homo sapien

543	25	89.3	661	2	AAB61883	Aab61883 homo sapi	616	24	85.7	76	2	Q8U856	Q8u856 agrobacteri
544	25	89.3	681	2	Q25346	Q25346 helicobacte	617	24	85.7	87	2	Q8D329	Q8d329 erwinia car
545	25	89.3	683	2	Q92LK7	Q92LK7 helicobacte	618	24	85.7	89	2	Q6LI08	Q6li08 bartonella
546	25	89.3	684	1	YF74_METUA	Y58969 methanococc	619	24	85.7	90	2	Q6YRL0	Q6yrl0 onion yello
547	25	89.3	689	2	Q6R7C3	Q6r7c3 ostrleid her	620	24	85.7	90	2	BAD04089	Bad04089 onion yel
548	25	89.3	689	2	AA500992	Aae00992 ostrleid h	621	24	85.7	91	2	Q48265	Q48265 haemophilus
549	25	89.3	705	2	Q6ZNC5	Q6znc5 homo sapien	622	24	85.7	96	2	Q832R9	Q832r9 enterococcu
550	25	89.3	705	2	BAD18450	Bad18450 homo sapi	623	24	85.7	97	2	Q7TU85	Q7tu85 prochloroco
551	25	89.3	724	2	Q9SCF0	Q9scf0 datura stra	624	24	85.7	106	2	Q8CG35	Q8cg35 rattus norv
552	25	89.3	729	2	Q6CC45	Q6cc45 yarrowia li	625	24	85.7	108	2	Q7Y3H2	Q7y3h2 enterobacte
553	25	89.3	742	2	Q7Q954	Q7q954 anopheles g	626	24	85.7	108	2	Q6FFA4	Q6ffa4 acinetobact
554	25	89.3	763	2	Q8I3P2	Q8i3p2 plasmodium	627	24	85.7	109	2	Q8KK54	Q8kk54 rhizobium e
555	25	89.3	774	2	Q7QH37	Q7qh37 anopheles g	628	24	85.7	116	2	Q9XAW0	Q9xaw0 pseudomonas
556	25	89.3	810	2	Q7MWJ0	Q7mwj0 porphyromon	629	24	85.7	119	2	Q7CTV2	Q7ctv2 agrobacteri
557	25	89.3	835	2	Q8BGK0	Q8bgk0 m mus muscu	630	24	85.7	133	1	RS19_ARCFU	RS19 archaesioglob
558	25	89.3	851	2	Q7Q1J5	Q7qlj5 anopheles g	631	24	85.7	135	2	Q66289	Q66289 pelargonium
559	25	89.3	865	1	TOP1_ECOLI	P06612 escherichia	632	24	85.7	135	2	Q84713	Q84713 pelargonium
560	25	89.3	865	1	TOP1_SALTY	P40686 salmonella	633	24	85.7	136	1	RS19_MEITH	RS19 methanobact
561	25	89.3	865	2	Q7AER8	Q7aek8 escherichia	634	24	85.7	139	2	Q6YRL1	Q6yrl1 onion yello
562	25	89.3	865	2	Q7UCR0	Q7ucr0 shigella fl	635	24	85.7	139	2	BAD04088	Bad04088 onion yel
563	25	89.3	865	2	Q8X7C5	Q8x7c5 escherichia	636	24	85.7	143	1	MR4Z_LACUO	MR4z lactobacilli
564	25	89.3	868	1	TOP1_HAEIN	P43012 haemophilus	637	24	85.7	143	1	MR4Z_STAEP	MR4z staphylococ
565	25	89.3	868	1	TOP1_PSEAE	Q9hzi5 pseudomonas	638	24	85.7	143	2	AAS08787	Aas08787 lactobaci
566	25	89.3	869	2	Q8ZB5	Q8zb5 pseudomonas	639	24	85.7	145	2	Q8B3U5	Q8b3u5 pelargonium
567	25	89.3	869	2	Q88K29	Q88k29 pseudomonas	640	24	85.7	150	2	Q9VUE3	Q9vue3 drosophila
568	25	89.3	871	2	Q6KAQ8	Q6kaq8 mus musculu	641	24	85.7	150	2	Q6HFU5	Q6hfs bacillus th
569	25	89.3	871	2	BAD21399	Bad21399 mus muscu	642	24	85.7	150	2	Q733N7	Q733n7 bacillus ce
570	25	89.3	879	2	Q83LC4	Q83lc4 shigella fl	643	24	85.7	150	2	Q81AG6	Q81ag6 bacillus ce
571	25	89.3	879	2	Q8FHU9	Q8fhu9 escherichia	644	24	85.7	150	2	AAS42526	Aas42526 bacillus
572	25	89.3	890	2	Q18196	Q18196 caenorhabdi	645	24	85.7	152	1	RS19_MEIJA	RS19 methanococc
573	25	89.3	912	1	ARH1_HUMAN	Q92888 homo sapien	646	24	85.7	152	2	Q8S022	Q8s022 oryza sativ
574	25	89.3	919	1	ARH1_RAT	Q9zli6 rattus norv	647	24	85.7	154	2	Q8L4J1	Q8l4j1 cucumis mel
575	25	89.3	920	1	ARH1_MOUSE	Q61210 mus musculu	648	24	85.7	154	2	Q8LK96	Q8lk96 cucumis mel
576	25	89.3	920	2	Q6YR75	Q6yri5 onion yello	649	24	85.7	155	2	Q9WV53	Q9wv53 cricetus
577	25	89.3	920	2	BAD04115	Bad04115 onion yel	650	24	85.7	160	2	Q7XYI8	Q7xyi8 chlorarachn
578	25	89.3	929	2	Q6KAM6	Q6kam6 mus musculu	651	24	85.7	161	2	Q71XZ2	Q71xz2 listeria mo
579	25	89.3	939	2	BAD21431	Bad21431 mus muscu	652	24	85.7	161	2	Q723I1	Q723i1 listeria mo
580	25	89.3	948	2	Q6NX52	Q6nx52 homo sapien	653	24	85.7	161	2	AAT03280	Aat03280 listeria
581	25	89.3	948	2	AAH67262	Aah67262 homo sapi	654	24	85.7	161	2	AAT04823	Aat04823 listeria
582	25	89.3	952	2	Q8K468	Q8k468 mus musculu	655	24	85.7	168	2	Q87846	Q87846 streptomyce
583	25	89.3	953	2	Q6ZNC7	Q6znc7 homo sapien	656	24	85.7	169	2	Q897V5	Q897v5 clostridium
584	25	89.3	953	2	BAD18448	Bad18448 homo sapi	657	24	85.7	171	2	Q7W2M0	Q7w2m0 bordetella
585	25	89.3	966	2	Q9VGX5	Q9vgx5 drosophila	658	24	85.7	171	2	Q7WDL6	Q7wdl6 bordetella
586	25	89.3	979	2	Q8INI2	Q8inl2 drosophila	659	24	85.7	179	2	Q9Y9G8	Q9y9g8 aeropyrum p
587	25	89.3	1000	2	P91550	P91550 caenorhabdi	660	24	85.7	180	2	Q9Z5A5	Q9z5a5 streptomyce
588	25	89.3	1032	2	Q7XSY1	Q7xsv1 oryza sativ	661	24	85.7	188	2	Q839V1	Q839v1 enterococu
589	25	89.3	1039	2	Q97Y47	Q97y47 sulfolobus	662	24	85.7	189	2	Q9HU76	Q9hu76 pseudomonas
590	25	89.3	1086	1	NTM_MOUSE	Q61941 mus musculu	663	24	85.7	197	2	Q6NAI5	Q6nai5 rhodospseud
591	25	89.3	1086	2	Q8C1W8	Q8ciw8 mus musculu	664	24	85.7	197	2	CAE28643	CAe28643 rhodopseu
592	25	89.3	1086	2	Q922E1	Q922e1 mus musculu	665	24	85.7	201	2	Q8NL60	Q8nl60 corynebacte
593	25	89.3	1088	2	Q9C9D7	Q9c9d7 arabidopsis	666	24	85.7	201	2	CAF13028	CAF13028 corynebac
594	25	89.3	1110	2	Q8GSY2	Q8gsy2 oryza sativ	667	24	85.7	204	2	Q7MBR9	Q7mbr9 vibrio vuln
595	25	89.3	1119	2	Q16715	Q16715 caenorhabdi	668	24	85.7	209	2	Q8N2W6	Q8n2w6 homo sapien
596	25	89.3	1260	2	Q9YCG9	Q9ycg9 aeropyrum p	669	24	85.7	213	2	Q9CPL7	Q9cpl7 pasteurella
597	25	89.3	1262	2	Q868L5	Q868l5 plasmodium	670	24	85.7	215	2	Q6HHN1	Q6hnn1 bacillus th
598	25	89.3	1262	2	Q8ID44	Q8id44 plasmodium	671	24	85.7	215	2	Q736K8	Q736k8 bacillus ce
599	25	89.3	1328	2	Q9YP35	Q9yp35 anabaena sp	672	24	85.7	215	2	Q81CA9	Q81ca9 bacillus ce
600	25	89.3	1330	2	Q55284	Q55284 synechocyst	673	24	85.7	215	2	AA541804	AA541804 bacillus
601	25	89.3	1331	2	P73020	P73020 synechocyst	674	24	85.7	217	2	Q9CG54	Q9cg54 lactococcus
602	25	89.3	1335	2	Q7VIN2	Q7vin2 prochloroco	675	24	85.7	219	2	Q8RGL8	Q8rgl8 fusobacteri
603	25	89.3	1336	2	Q7TTW0	Q7ttw0 synechococc	676	24	85.7	224	1	DLTR_STR43	DLTR streptococc
604	25	89.3	1337	2	Q7VC39	Q7vc39 prochloroco	677	24	85.7	225	2	Q6T1F0	Q6t1f0 sars corona
605	25	89.3	1347	2	Q7TU23	Q7tu23 prochloroco	678	24	85.7	225	2	AAS01056	Aas01056 sars coro
606	25	89.3	1433	2	Q8DLX7	Q8dlx7 synechococc	679	24	85.7	229	2	Q7P8F1	Q7p8f1 fusobacteri
607	25	89.3	1483	2	Q8YIC4	Q8yyc4 anabaena sp	680	24	85.7	240	2	Q8TX01	Q8tx01 methanopyru
608	25	89.3	2413	1	PRP8_YEAST	P33334 saccharomyc	681	24	85.7	244	2	Q7MYR5	Q7myr5 photorhabdu
609	25	89.3	3166	2	Q9W3Z0	Q9w3z0 drosophila	682	24	85.7	244	2	Q8YK53	Q8yk53 anabaena sp
610	24	85.7	50	2	Q9TUX6	Q9tux6 canis famil	683	24	85.7	250	2	Q8A412	Q8a412 bacteroides
611	24	85.7	51	2	Q8LLV4	Q8llv4 ipomea bat	684	24	85.7	251	2	Q9UYK8	Q9uyk8 pyrococcus
612	24	85.7	67	2	Q72DE2	Q72de2 desulfovibr	685	24	85.7	253	2	Q6JBI2	Q6jbi2 dictyocaulu
613	24	85.7	67	2	AA595467	Aae595467 desulfovi	686	24	85.7	253	2	AAT06311	Aat06311 dictyocau
614	24	85.7	72	2	Q9KXC5	Q9kcx5 bacillus ha	687	24	85.7	255	2	Q92WM8	Q92wm8 oryza sativ
615	24	85.7	74	1	YMUC_SALTY	P14500 salmonella	688	24	85.7	256	2	Q38150	Q38150 bacterioph

689	24	85.7	256	2	Q72YK0	Q72YK0 bacillus ce	762	24	85.7	395	1	UXUA_VIBPA	Q87fh9 vibrio para
690	24	85.7	256	2	AA543922	AA543922 bacillus	763	24	85.7	395	1	UXUA_VIBVU	Q8d562 vibrio vuln
691	24	85.7	258	2	Q7ID52	Q7d522 halovirus h	764	24	85.7	395	1	UXUA_VIBVY	Q7mbz9 vibrio vuln
692	24	85.7	258	2	Q8V6J6	Q8v6j6 halovirus h	765	24	85.7	398	2	Q93FW0	Q93fw0 ehrlichia s
693	24	85.7	259	2	Q9SN00	Q9snq0 oryza sativ	766	24	85.7	398	2	Q93FW1	Q93fw1 ehrlichia r
694	24	85.7	260	2	Q6YQ47	Q6yq47 onion yello	767	24	85.7	399	1	ACKA_SHEON	Q8ed55 shewanella
695	24	85.7	260	2	BAD04613	BAD04613 onion yel	768	24	85.7	401	2	Q8FV93	Q8fv93 bruceella su
696	24	85.7	262	2	Q81240	Q81240 arabidopsis	769	24	85.7	401	2	Q92LQ8	Q92lq8 rhizobium m
697	24	85.7	262	2	Q745D3	Q745d3 mycobacteri	770	24	85.7	401	2	Q8YD33	Q8yd33 bruceella me
698	24	85.7	262	2	AA502461	AA502461 mycobacte	771	24	85.7	402	2	Q8YD42	Q8y42 agrobacteri
699	24	85.7	263	2	Q8LCD9	Q8lcd9 arabidopsis	772	24	85.7	402	2	Q8U8G9	Q8u8g9 agrobacteri
700	24	85.7	263	2	Q9FMX0	Q9fmx0 arabidopsis	773	24	85.7	418	2	Q8X4H7	Q8x4h7 escherichia
701	24	85.7	263	2	Q9Z4G8	Q9z4g8 plasmid pkm	774	24	85.7	427	1	Y867 METJA	Q58277 methanococ
702	24	85.7	263	2	Q798C1	Q79sc1 incn plasm	775	24	85.7	428	1	APG2 METJA	Q60326 methanococ
703	24	85.7	264	2	Q881G3	Q881g3 staphylococ	776	24	85.7	437	2	Q873M9	Q873m9 mucor circi
704	24	85.7	268	2	Q826D4	Q826d4 streptomyce	777	24	85.7	462	2	Q8RAI2	Q8rai2 thermoanaer
705	24	85.7	270	2	Q6HN82	Q6hn82 bacillus th	778	24	85.7	472	2	Q9X1X5	Q9x1x5 thermotoga
706	24	85.7	270	2	Q81UX2	Q81ux2 bacillus an	779	24	85.7	488	2	YAHG_ECOLI	Y77221 escherichia
707	24	85.7	270	2	AA229842	AA229842 bacillus	780	24	85.7	492	2	Q8MQU6	Q8mq6 dictyoscell
708	24	85.7	272	2	Q8SVD1	Q8svd1 encephalito	781	24	85.7	492	2	Q8MQU6	Q8mq6 dictyoscell
709	24	85.7	274	2	Q8LHN2	Q8lhn2 arabidopsis	782	24	85.7	495	2	Q7UPW5	Q7upw5 rhodospirell
710	24	85.7	292	2	Q7X347	Q7x347 uncultured	783	24	85.7	500	2	Q6EUF8	Q6euf8 oryza sativ
711	24	85.7	292	2	Q6UNH5	Q6unh5 ictalurus p	784	24	85.7	504	1	GALL_CANPA	O42821 candida par
712	24	85.7	292	2	AAQ72816	AAq72816 ictalurus	785	24	85.7	504	1	GALL_CANPA	O42821 candida par
713	24	85.7	296	1	ARGI_BACSU	P39138 bacillus su	786	24	85.7	516	1	CYS2_CUCMA	Q91x5 arabidopsis
714	24	85.7	296	2	Q933D8	Q933q8 clostridium	787	24	85.7	519	2	Q7MEU1	Q7meu1 vibrio vuln
715	24	85.7	300	2	Q72IK0	Q72ik0 thermus the	788	24	85.7	519	2	Q8D7S7	Q8d7s7 vibrio vuln
716	24	85.7	300	2	AA581474	AA581474 thermus t	789	24	85.7	522	1	COR1_SCHPO	O13688 schizosacch
717	24	85.7	304	1	Y191_METHH	Q6293 methanobact	790	24	85.7	522	1	Q839T7	Q839t7 enterococcu
718	24	85.7	305	2	Q6Z6U8	Q6z6u8 oryza sativ	791	24	85.7	522	2	Q8PYL3	Q8pyl3 methanosarc
719	24	85.7	305	2	BAD15947	BAD15947 oryza sat	792	24	85.7	523	2	Q8THZ6	Q8thz6 methanosarc
720	24	85.7	310	1	RNZ_TREPA	O07896 treponema p	793	24	85.7	528	2	Q6BIZ6	Q6biz6 debaromyce
721	24	85.7	313	2	Q8TWH1	Q8twh1 methanopyru	794	24	85.7	531	2	Q874U9	Q874u9 emericella
722	24	85.7	314	1	Y057_EUCAP	Q8ka54 buchnera ap	795	24	85.7	534	2	Q8ZUM2	Q8zuw2 pyrobaculum
723	24	85.7	315	1	Y060_BUCAI	P57168 buchnera ap	796	24	85.7	534	1	SYM_THEVO	Q82uw2 pyrobaculum
724	24	85.7	320	2	Q7QMA7	Q7qma7 anopheles g	797	24	85.7	543	1	HUTU_BACAN	Q979b7 thermoplas
725	24	85.7	325	2	Q8R875	Q8r875 thermoanaer	798	24	85.7	543	1	SYM_THEVO	Q979b7 thermoplas
726	24	85.7	334	1	NDA_VERPE	O8xgm5 versinia pe	799	24	85.7	552	2	AA542584	AA542584 bacillus
727	24	85.7	342	2	Q8XJ13	O8xj13 clostridium	800	24	85.7	552	2	AA542584	AA542584 bacillus
728	24	85.7	346	2	Q6JBI3	Q6jbi3 dictyocaulu	801	24	85.7	552	2	AA542584	AA542584 bacillus
729	24	85.7	346	2	Q987Q0	Q987q0 rhizobium l	802	24	85.7	552	2	AA542584	AA542584 bacillus
730	24	85.7	346	2	AA542584	AA542584 dictyocau	803	24	85.7	552	2	AA542584	AA542584 bacillus
731	24	85.7	349	1	GBT1_BOVIN	P04695 bos taurus	804	24	85.7	552	2	AA542584	AA542584 bacillus
732	24	85.7	349	1	GBT1_CANPA	Q28300 canis fami	805	24	85.7	552	2	AA542584	AA542584 bacillus
733	24	85.7	349	1	GBT1_HUMAN	P11488 homo sapien	806	24	85.7	553	2	Q8X3S6	Q8x3s6 escherichia
734	24	85.7	349	1	GBT1_MOUSE	P20612 mus musculu	807	24	85.7	553	2	Q8X3S6	Q8x3s6 escherichia
735	24	85.7	350	2	Q9DG28	Q9dg28 gallus gall	808	24	85.7	557	1	HRBA_AQUAE	Q52498 aquifex aeo
736	24	85.7	350	2	Q9DXX6	Q9dwx6 brachydanio	809	24	85.7	557	1	HRBA_AQUAE	Q52498 aquifex aeo
737	24	85.7	350	2	Q6DJF0	Q6djf0 xenopus lae	810	24	85.7	568	1	HRBA_AQUAE	Q52498 aquifex aeo
738	24	85.7	350	2	CAD45009	CAD45009 canis fam	811	24	85.7	570	2	AA542584	AA542584 bacillus
739	24	85.7	350	2	AA545946	AA545946 brachydan	812	24	85.7	585	2	AA542584	AA542584 bacillus
740	24	85.7	351	1	CARA_CLOAB	Q97ft2 clostridium	813	24	85.7	585	2	AA542584	AA542584 bacillus
741	24	85.7	352	2	Q8TWH9	Q8twh9 methanopyru	814	24	85.7	592	1	SVFB_CANAL	Q8tpv4 methanosarc
742	24	85.7	352	2	Q942L6	O942l6 sclerotinia	815	24	85.7	592	1	SVFB_CANAL	Q8tpv4 methanosarc
743	24	85.7	359	2	Q834E1	Q834e1 enterococcu	816	24	85.7	593	2	Q6C4X4	Q6c4x4 candida alb
744	24	85.7	361	2	Q9A2C3	Q9a2c3 caulobacter	817	24	85.7	594	1	SVFB_YEAST	P15624 saccharomyc
745	24	85.7	366	1	MINC_ANASP	O8yrj1 anabaena sp	818	24	85.7	599	1	CENB_MOUSE	P27790 mus musculu
746	24	85.7	366	2	Q72YL8	Q72yl8 bacillus ce	819	24	85.7	599	2	AAH71269	AAh71269 mus muscu
747	24	85.7	366	2	AA543904	AA543904 bacillus	820	24	85.7	606	1	CENB_CRIGR	P48988 cricetulus
748	24	85.7	369	2	Q8ELL8	Q8ell8 oceanobacil	821	24	85.7	607	2	Q8Q077	Q8q077 methanosarc
749	24	85.7	370	1	YB11_HALN1	P17103 halobacteri	822	24	85.7	608	2	Q6CEX5	O6cex5 varrowia li
750	24	85.7	373	1	Y007_METJA	Q60318 methanococc	823	24	85.7	608	2	Q75BQ2	Q75bq2 ashbya gos
751	24	85.7	378	2	Q89VE1	Q89ve1 bradyrhizob	824	24	85.7	609	2	AA551445	AA551445 ashbya go
752	24	85.7	378	2	Q8YUR9	Q8yur9 anabaena sp	825	24	85.7	611	2	Q9S767	Q9s767 oryza sativ
753	24	85.7	381	1	PRC1_CORGL	Q8nsh7 corynebacte	826	24	85.7	612	2	Q73RP0	Q73rp0 treponema d
754	24	85.7	383	1	PRC2_CORGL	Q8nsh7 corynebacte	827	24	85.7	614	2	AA510543	AA510543 treponema
755	24	85.7	384	1	CISY_THEAC	P21553 thermoplas	828	24	85.7	614	2	AA510543	AA510543 treponema
756	24	85.7	385	2	Q595F69	Q595f69 methanobact	829	24	85.7	622	2	Q923C5	Q923c5 mus musculu
757	24	85.7	386	2	Q97C64	Q97c64 thermoplas	830	24	85.7	622	2	Q923C5	Q923c5 mus musculu
758	24	85.7	387	2	O27638	O27638 methanobact	831	24	85.7	632	1	GAAT_HUMAN	Q99874 glycine max
759	24	85.7	389	2	Q72W91	Q72w91 leptospira	832	24	85.7	632	1	GAAT_HUMAN	Q99874 glycine max
760	24	85.7	389	2	Q8F9Y5	Q8f9y5 leptospira	833	24	85.7	632	1	GAAT_HUMAN	Q99874 glycine max
761	24	85.7	389	2	AA568683	AA568683 leptospir	834	24	85.7	637	2	Q9A5B9	Q9a5e9 caulobacter

835	24	85.7	638	1	GAAT_MOUSE	Q9jlf1 mus musculus	908	24	85.7	1154	2	Q89Tl2	Q89t12 bradyrhizob
836	24	85.7	692	1	QNM618	Q9m618 brassica ju	909	24	85.7	1175	2	Q98L07	Q98l07 rhizobium l
837	24	85.7	701	1	ATB1_ANASP	Q8ype9 anabaena sp	910	24	85.7	1191	2	Q72T01	Q72t01 leptospira
838	24	85.7	712	2	Q3ZLE7	Q9zle7 helicobacte	911	24	85.7	1191	2	Q8F2G6	Q8f2g6 leptospira
839	24	85.7	713	2	Q25402	Q25402 helicobacte	912	24	85.7	1191	2	Q8A69827	Q8a69827 leptospir
840	24	85.7	713	2	Q7MRX3	Q7mrx3 wolinnella s	913	24	85.7	1244	2	Q55576	Q55576 synechocyst
841	24	85.7	714	2	Q8C0U1	Q8c0u1 mus musculus	914	24	85.7	1328	1	FINC_PLEWA	Q91289 pleurodeles
842	24	85.7	715	2	Q7F7Q2	Q7p7q2 fusobacteri	915	24	85.7	1343	2	Q6CY12	Q6cy12 kluyveromyc
843	24	85.7	724	2	Q9HN37	Q9hn37 halobacteri	916	24	85.7	1348	2	Q8TT85	Q8tt85 methanosarc
844	24	85.7	725	1	SPE1_DIACA	Q96412 dianthus ca	917	24	85.7	1428	2	Q8TR62	Q8tr62 methanosarc
845	24	85.7	725	2	Q04429	Q04429 dianthus ca	918	24	85.7	1443	2	Q8MTB2	Q8mtb2 drosophila
846	24	85.7	749	2	Q7T4K3	Q7t4k3 pea stem ne	919	24	85.7	1555	2	Q7PPH8	Q7pph8 anopheles g
847	24	85.7	756	2	Q70WU1	Q70wu1 pelargonium	920	24	85.7	1562	2	Q9Y840	Q9y840 mycosphaere
848	24	85.7	756	2	CAD55835	Cad55835 petargoni	921	24	85.7	1584	2	Q8Q0G8	Q8q0g8 methanosarc
849	24	85.7	760	2	Q8R6F6	Q8r6f6 fusobacteri	922	24	85.7	1765	2	Q9VS30	Q9vs30 drosophila
850	24	85.7	762	2	Q9IRAI	Q9irai japanese ir	923	24	85.7	1770	2	Q9VS29	Q9vs29 drosophila
851	24	85.7	763	2	O6EN28	O6en28 pelargonium	924	24	85.7	1908	2	Q7MT05	Q7mt05 porphyromon
852	24	85.7	774	2	Q8S3P4	Q8s3p4 oryza sativ	925	24	85.7	1957	2	Q91YC9	Q91yc9 mus musculu
853	24	85.7	774	2	CAE02886	Ca02886 oryza sat	926	24	85.7	1972	1	P531_HUMAN	Q12888 homo sapien
854	24	85.7	774	2	CAE54575	Ca54575 oryza sat	927	24	85.7	1977	2	Q7Z3U4	Q7z3u4 homo sapien
855	24	85.7	791	2	Q893L2	Q893l2 clostridium	928	24	85.7	2015	2	Q7QRQ0	Q7qrq0 giardia lam
856	24	85.7	791	2	Q6VRG4	Q6vrg4 melon necro	929	24	85.7	2258	2	Q7QF47	Q7qf47 anopheles g
857	24	85.7	791	2	Q83424	Q83424 melon necro	930	24	85.7	2265	1	FINC_BOVIN	P07589 bos taurus
858	24	85.7	791	2	Q8JMI2	Q8jmi2 melon necro	931	24	85.7	4382	2	Q6JH39	Q6jh39 sars corona
859	24	85.7	791	2	Q8V992	Q8v992 melon necro	932	24	85.7	4382	2	Q6JH47	Q6jh47 sars corona
860	24	85.7	791	2	Q8V996	Q8v996 melon necro	933	24	85.7	4382	2	Q6R7Y7	Q6r7y7 sars corona
861	24	85.7	791	2	AA01965	AA01965 melon nec	934	24	85.7	4382	2	Q6RCW6	Q6rcw6 sars corona
862	24	85.7	793	2	Q6ML73	Q6ml73 bdellovibri	935	24	85.7	4382	2	Q6RCX7	Q6rcx7 sars corona
863	24	85.7	793	2	CAE79984	Ca79984 bdellovib	936	24	85.7	4382	2	Q6RCY8	Q6rcy8 sars corona
864	24	85.7	794	2	Q8GYI8	Q8gyi8 caenorhabdi	937	24	85.7	4382	2	Q6RCZ9	Q6rcz9 sars corona
865	24	85.7	823	2	Q6FBES	Q6fbes acinetobact	938	24	85.7	4382	2	Q6RD10	Q6rd10 sars corona
866	24	85.7	828	2	Q6RX39	Q6rx39 arabidopsis	939	24	85.7	4382	2	Q6RD21	Q6rd21 sars corona
867	24	85.7	828	2	AA593962	AA593962 arabidopsi	940	24	85.7	4382	2	Q6RD32	Q6rd32 sars corona
868	24	85.7	832	2	Q8RF51	Q8rf51 fusobacteri	941	24	85.7	4382	2	Q6RD43	Q6rd43 sars corona
869	24	85.7	838	2	Q8FG37	Q8fg37 arabidopsis	942	24	85.7	4382	2	Q6RD54	Q6rd54 sars corona
870	24	85.7	838	2	AAQ65168	Aaq65168 arabidops	943	24	85.7	4382	2	Q6RD55	Q6rd55 sars corona
871	24	85.7	839	2	Q9UT23	Q9ut23 schizosacch	944	24	85.7	4382	2	Q6VA79	Q6va79 sars corona
872	24	85.7	842	2	Q7KIS3	Q7kis3 mycoplasma	945	24	85.7	4382	2	Q6VAA1	Q6vaa1 sars corona
873	24	85.7	842	2	AA272502	AA272502 mycoplas	946	24	85.7	4382	2	Q6VAA9	Q6vaa9 sars corona
874	24	85.7	852	2	Q8K940	Q8k940 bacillus ha	947	24	85.7	4382	2	Q6WGN0	Q6wgn0 sars corona
875	24	85.7	865	1	LOX2_SOYEN	P09439 glycine max	948	24	85.7	4382	2	Q7T697	Q7t697 sars corona
876	24	85.7	866	2	Q39870	Q39870 glycine max	949	24	85.7	4382	2	AA91585	AA91585 sars coro
877	24	85.7	871	2	Q70WU2	Q70wu2 pelargonium	950	24	85.7	4382	2	AA16181	AA16181 sars coro
878	24	85.7	871	2	CAD55834	Cad55834 pelargoni	951	24	85.7	4382	2	AA94736	AA94736 sars coro
879	24	85.7	886	2	Q8TOA5	Q8tqa5 methanosarc	952	24	85.7	4382	2	AA94747	AA94747 sars coro
880	24	85.7	888	2	Q8IRA2	Q8ira2 japanese ir	953	24	85.7	4382	2	AA94758	AA94758 sars coro
881	24	85.7	904	2	Q8AIK1	Q8aik1 bacteroides	954	24	85.7	4382	2	AA87500	AA87500 sars coro
882	24	85.7	907	2	Q72JY9	Q72jy9 thermus the	955	24	85.7	4382	2	AA87511	AA87511 sars coro
883	24	85.7	907	2	AA80977	AA80977 thermus t	956	24	85.7	4382	2	AA87522	AA87522 sars coro
884	24	85.7	911	2	Q8UVU0	Q8uvu0 xenopus lae	957	24	85.7	4382	2	AA87533	AA87533 sars coro
885	24	85.7	954	2	Q83EP1	Q83ep1 coxiella bu	958	24	85.7	4382	2	AA87544	AA87544 sars coro
886	24	85.7	964	1	RPO_MCMV	P11640 maize chlor	959	24	85.7	4382	2	AA87555	AA87555 sars coro
887	24	85.7	965	2	Q8IBM2	Q8ibm2 maize chlor	960	24	85.7	4382	2	AA87566	AA87566 sars coro
888	24	85.7	1014	2	Q8CI97	Q8ci97 mus musculu	961	24	85.7	4382	2	AA87577	AA87577 sars coro
889	24	85.7	1027	2	Q41930	Q41930 murid herpe	962	24	85.7	4382	2	AA87588	AA87588 sars coro
890	24	85.7	1039	2	Q73QS7	Q73qs7 treponema d	963	24	85.7	4382	2	AA87599	AA87599 sars coro
891	24	85.7	1039	2	AA510861	Aa510861 treponema	964	24	85.7	6880	2	Q6S8D8	Q6s8d8 sars corona
892	24	85.7	1047	2	Q8A8Z1	Q8a8z1 caulobacter	965	24	85.7	6880	2	AA86774	AA86774 sars coro
893	24	85.7	1050	2	Q97XX8	Q97xx8 sulfolobus	966	24	85.7	7073	1	RIAB_CVISA	P59641 h replicase
894	24	85.7	1062	2	Q95Q61	Q95q61 caenorhabdi	967	24	85.7	7073	2	Q6JH40	Q6jh40 sars corona
895	24	85.7	1063	2	Q7PMK4	Q7pmk4 anopheles g	968	24	85.7	7073	2	Q6JH48	Q6jh48 sars corona
896	24	85.7	1064	1	SVI_ARPE	Q9y6f7 aeropyrum p	969	24	85.7	7073	2	Q6RWY8	Q6rwy8 sars corona
897	24	85.7	1066	2	Q8TQSO	Q8tgs0 methanosarc	970	24	85.7	7073	2	Q6RCW7	Q6rcw7 sars corona
898	24	85.7	1071	1	CARB_BACSU	P25994 bacillus su	971	24	85.7	7073	2	Q6RCX8	Q6rcx8 sars corona
899	24	85.7	1091	2	Q8Y4J2	Q8y4j2 listeria mo	972	24	85.7	7073	2	Q6RCY9	Q6rcy9 sars corona
900	24	85.7	1111	2	Q86FP2	Q86fp2 caenorhabdi	973	24	85.7	7073	2	Q6RD00	Q6rd00 sars corona
901	24	85.7	1111	2	Q73I04	Q73i04 wolbachia p	974	24	85.7	7073	2	Q6RD11	Q6rd11 sars corona
902	24	85.7	1111	2	AA514108	Aa514108 wolbachia	975	24	85.7	7073	2	Q6RD22	Q6rd22 sars corona
903	24	85.7	1116	2	Q7CW80	Q7cw80 agrobacteri	976	24	85.7	7073	2	Q6RD33	Q6rd33 sars corona
904	24	85.7	1116	2	Q8UBP7	Q8ubp7 agrobacteri	977	24	85.7	7073	2	Q6RD44	Q6rd44 sars corona
905	24	85.7	1121	2	Q8W302	Q8w302 oryza sativ	978	24	85.7	7073	2	Q6RD55	Q6rd55 sars corona
906	24	85.7	1126	2	Q8TKN1	Q8tkn1 methanosarc	979	24	85.7	7073	2	Q6RD66	Q6rd66 sars corona
907	24	85.7	1127	2	Q9N323	Q9n323 caenorhabdi	980	24	85.7	7073	2	Q6TPE9	Q6tpe9 sars corona

```
981 24 85.7 7073 2 Q6UZF1 Q6uzf1 sars corona
982 24 85.7 7073 2 Q6UZF5 Q6uzf5 sars corona
983 24 85.7 7073 2 Q6V586 Q6v586 sars corona
984 24 85.7 7073 2 Q6VA80 Q6va80 sars corona
985 24 85.7 7073 2 Q6VA91 Q6va91 sars corona
986 24 85.7 7073 2 Q6VAA2 Q6vaa2 sars corona
987 24 85.7 7073 2 AAR14802 Aar14802 sars coro
988 24 85.7 7073 2 AAR14806 Aar14806 sars coro
989 24 85.7 7073 2 AAR14810 Aar14810 sars coro
990 24 85.7 7073 2 AAP82978 Aap82978 sars coro
991 24 85.7 7073 2 AAR91584 Aar91584 sars coro
992 24 85.7 7073 2 AAS00002 Aas00002 sars coro
993 24 85.7 7073 2 AAP49011 Aap49011 sars coro
994 24 85.7 7073 2 AAP94735 Aap94735 sars coro
995 24 85.7 7073 2 AAP94746 Aap94746 sars coro
996 24 85.7 7073 2 AAP94757 Aap94757 sars coro
997 24 85.7 7073 2 AAR87499 Aar87499 sars coro
998 24 85.7 7073 2 AAR87510 Aar87510 sars coro
999 24 85.7 7073 2 AAR87521 Aar87521 sars coro
1000 24 85.7 7073 2 AAR87532 Aar87532 sars coro
```

## ALIGNMENTS

```
RESULT 1
Q6D6P6 PRELIMINARY; PRT; 78 AA.
AC Q6D6P6;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN ORNames=ECAL636;
OS Erwinia carotovora subsp. atroseptica SCRI1043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1043;
RA Bell K.S., Sebahia M., Pritchard L., Holden M., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.;
RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BX950851; CAG74540.1; -.
KW Hypothetical protein.
SQ SEQUENCE 78 AA; 8646 MW; 52B7EC25B09B9DD9 CRC64;
```

```
Query Match 96.4%; Score 27; DB 2; Length 78;
Best Local Similarity 83.3%; Pred. NO. 97;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 GYXVEE 6
DB 35 GYSVEE 40
|||
```

```
RESULT 2
Q982Z2 PRELIMINARY; PRT; 94 AA.
AC Q982Z2;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transcriptional regulator.
GN OrderedLocusNames=mlr8432;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
```

```
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Igesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082936; PubMed=11214974;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Igesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti (supplement).";
RL DNA Res. 7:381-406(2000).
DR EMBL; AP003014; BAB54314.1; -.
GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH 3.
DR InterPro; IPR010982; Lambda_like_DNA.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
DR PROSITE; PS50943; HTH_CROCI; 1.
KW Complete proteome.
SQ SEQUENCE 94 AA; 9895 MW; 02CD173EB020CF25 CRC64;
```

```
Query Match 96.4%; Score 27; DB 2; Length 94;
Best Local Similarity 83.3%; Pred. NO. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 GYXVEE 6
DB 33 GYSVEE 38
|||
```

```
RESULT 3
Q9AAK1 PRELIMINARY; PRT; 122 AA.
ID Q9AAK1
AC Q9AAK1; 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chemotaxis protein CheYIV.
GN OrderedLocusNames=CC0596;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Utterback T.R., Iran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005734; AAK22582.1; -.
DR FIR; B87323; B87323.
```

```

DR HSP; P06143; 1JBE.
DR TIGR; CC0596; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR011006; CheY_like.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00072; Response_reg.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS01110; RESPONSE_REGULATORY; 1.
KW Complete proteome; Phosphorylation; Sensory transduction.
SQ SEQUENCE 122 AA; 13361 MW; A48FAD9D9886B7DD CRC64;

Query Match 96.4%; Score 27; DB 2; Length 122;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 25 GYAVEE 30

RESULT 4
Q6E432 PRELIMINARY; PRT; 167 AA.
AC Q6E432;
DT 01-OCT-2004 (TRENBLrel. 28, Created)
DT 01-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE RNA-dependent RNA polymerase (fragment).
OS Nootka lupine vein-clearing virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae.
OX NCBI_TaxID=283876;
RN [1]
RP SEQUENCE FROM N.A.
RA Robertson N.L.;
RT "The biology of a new virus isolated from Lupinus nootkatensis plants in Alaska.";
RL Plant Pathol. 0:0-0(2004).
DR EMBL; AY584590; AAT69242.1; -.
KW RNA-directed RNA polymerase.
FT NON TER 1
FT NON TER 167
SQ SEQUENCE 167 AA; 18662 MW; D3D14C61086A9707 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 167;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 44 GYVEE 49

RESULT 5
Q897V6 PRELIMINARY; PRT; 168 AA.
ID Q897V6;
AC Q897V6;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE RNA polymerase sigma factor.
OS OrderedLocusNames=CTC00618;
ON Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88;
RX MEDLINE=22457253; PubMed=12552129;

```

```

RA Brueggemann H., Baumer S., Fricke W.F., Wierzer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gotschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR EMBL; AE015938; AAC35230.1; -.
DR HSP; P34086; 1OR7.
DR GO; GO:0016987; F:sigma factor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006352; P:transcription initiation; IEA.
DR InterPro; IPR009059; bi_resp_regltr_C.
DR InterPro; IPR009043; RNA_pol_sigma.
DR InterPro; IPR007627; Sigma70_r2.
DR InterPro; IPR007630; Sigma70_r4.
DR Pfam; PF04542; Sigma70_r2; 1.
DR Pfam; PF04545; Sigma70_r4; 1.
KW Complete proteome.
SQ SEQUENCE 168 AA; 20061 MW; 827C5B6C882E3C51 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 168;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 128 GYSVEE 133

RESULT 6
Q8ZYH3 PRELIMINARY; PRT; 170 AA.
ID Q8ZYH3;
AC Q8ZYH3;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein PAE0776.
OS OrderedLocusNames=PAE0776;
OC Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009785; AAL63020.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 170 AA; 18494 MW; 57189DE94397798E CRC64;

Query Match 96.4%; Score 27; DB 2; Length 170;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 129 GYAVEE 134

RESULT 7
Q7ZUZO PRELIMINARY; PRT; 191 AA.
ID Q7ZUZO;
AC Q7ZUZO;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Similar to Cof9 homolog (Hypothetical protein).

```

GN Name=29c:56319;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo, and Whole body;  
 RX MEDLINE=2238257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stاپleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole body;  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Strausberg R.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC046071; AAH46071.1; -;  
 DR EMBL; BC071487; AAH71487.1; -;  
 DR InterPro; IPR000717; PCI.  
 KW Hypothetical protein.  
 SQ SEQUENCE 191 AA; 21546 MW; 7E36460743586020 CRC64;  
  
 Query Match 96.4%; Score 27; DB 2; Length 191;  
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 GYXVEE 6  
 DB 136 GYSVEE 141  
  
 RESULT 8  
 AAH71487  
 ID AAH71487 PRELIMINARY; PRT; 191 AA.  
 AC AAH71487;  
 DT 01-JUN-2004 (TrEMBLrel. 27, Created)  
 DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=2238257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stاپleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Strausberg R.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC071487; AAH71487.1; -;  
 DR EMBL; BC071487; AAH71487.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 191 AA; 21546 MW; 7E36460743586020 CRC64;  
  
 Query Match 96.4%; Score 27; DB 2; Length 191;  
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 GYXVEE 6  
 DB 136 GYSVEE 141  
  
 RESULT 9  
 Q52048  
 ID Q52048 PRELIMINARY; PRT; 200 AA.  
 AC Q52048;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS plasmid phv2.  
 OG plasmid phv2.  
 OC plasmids.  
 OX NCBI\_TaxID=2562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88068622; PubMed=2825193;  
 RA Charlebois R.L., Lam W.L., Cline S.W., Doolittle W.F.;  
 RT "Characterization of phv2 from Halobacterium volcanii and its use in  
 RL demonstrating transformation of an archaeobacterium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8530-8534 (1987).  
 DR EMBL; J03014; AAA88100.1; -;  
 KW Hypothetical protein; plasmid.  
 SQ SEQUENCE 200 AA; 22918 MW; C86BF16C7A57DB96 CRC64;  
  
 Query Match 96.4%; Score 27; DB 2; Length 200;  
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 GYXVEE 6  
 DB 119 GYSVEE 124  
  
 RESULT 10  
 Q7N5F1  
 ID Q7N5F1 PRELIMINARY; PRT; 231 AA.  
 AC Q7N5F1;



```

DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Similar to probable transcription regulator.
GN OrderedLocNames=plu2001;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Rusnok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taouit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Siguier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
RL EMBL; BX571865; CAE14294.1; -.
DR PhotoList; plu2001; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; E:transcription factor activity; IEA.
DR GO; GO:0006355; F:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000792; HTH_LuxR.
KW Complete proteome.
SQ SEQUENCE 231 AA; 26996 MW; 845723D1DF51125B CRC64;

Query Match 96.4%; Score 27; DB 2; Length 231;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 57 GYTVEE 52

RESULT 11
Q7XYK9 PRELIMINARY; PRT; 232 AA.
AC Q7XYK9;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Plastid protein SufE (Fragment).
OS Chlorarachnion sp. (strain CCMP 621) (Pedinomonas minutissima).
OC Eukaryota; Cercozoa; Chlorarachniophyceae; Bigelowiella.
OX NCBI_TaxID=227086;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCMP 621;
RX MEDLINE=22709102; PubMed=12777624;
RA Archibald J.M., Rogers M.B., Toop M., Ishida K., Keeling P.J.;
RT "Lateral gene transfer and the evolution of plastid-targeted proteins
RT in the secondary plastid-containing alga Bigelowiella natans.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7678-7683(2003).
DR EMBL; AY267671; AAP79185.1; -.
DR InterPro; IPR003808; SufE.
DR Pfam; PF02657; SufE; 1.
FT NON TER 1
SQ SEQUENCE 232 AA; 25504 MW; 75207222030C6428 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 232;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 184 GYSVEE 189

RESULT 12
Q9UZZ8 PRELIMINARY; PRT; 233 AA.
AC Q9UZZ8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=PYRAB10000; ORFNames=PAB1706;
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RX MEDLINE=22511545; PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
RA Poch O., Frier D., Querellou J., Ripp R., Thierry J.-C.,
RA van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RT archaeon Pyrococcus abyssi.";
RL Mol. Microbiol. 47:1495-1512(2003).
DR EMBL; AJ248286; CAB49908.1; -.
DR PIR; G75075; G75075.
DR InterPro; IPR002826; DUF115.
DR InterPro; IPR000886; ER_target_S.
DR Pfam; PF01973; DUF115; 1.
DR ProDom; PD016917; DUF115; 1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 233 AA; 26421 MW; 586316A291510C2F CRC64;

Query Match 96.4%; Score 27; DB 2; Length 233;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 18 GYSVEE 23

RESULT 13
Q9KJX2 PRELIMINARY; PRT; 233 AA.
AC Q9KJX2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE PhoP3 response regulator.
GN Name=phoP3;
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytophacteriineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DZF1;
RX MEDLINE=22450566; PubMed=12562808;
RA Moraleda-Munoz A., Carrero-Lerida J., Perez J., Munoz-Dorado J.;
RT "Role of Two Novel Two-Component Regulatory Systems in Development and
RT Phosphatase Expression in Myxococcus xanthus.";
RL J. Bacteriol. 185:1376-1383(2003).
CC -!- SIMILARITY: Contains 1 response regulatory domain.
DR EMBL; AF157829; AAF82619.1; -.
DR HSPF; P08402; 1800.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR011006; CheY_like.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.

```

```
DR Pfam: PF00072; Response reg; 1.
DR Pfam: PF00486; Trans_reg_C; 1.
DR ProDom: PD000039; Response_reg; 1.
DR ProDom: PD000329; Trans_reg_C; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS01110; RESPONSE REGULATORY; 1.
KW DNA-binding; Phosphorylation; Sensory transduction;
KW Transcription regulation.
SQ SEQUENCE 233 AA; 26189 MW; 689C31D678C1F55 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 233;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 30 GYAVEE 35

RESULT 14
VG85_BPPH2
ID VG85_BPPH2 STANDARD; PRT; 280 AA.
AC P20344;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Head fiber protein (late protein GP8.5).
GN Name=8.5;
OS Bacteriophage phi-29.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC phi-29-like viruses.
OC NCBI_TaxID=10756;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87106857; PubMed=3803926;
RA Vicek C., Paces V.;
RT Nucleotide sequence of the late region of Bacillus phage phi 29
RT completes the 19,285-bp sequence of phi 29 genome. Comparison with the
RT homologous sequence of phage PZA.;
RL Gene 46:215-225(1986).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14782; AAA32281.1; -.
DR PIR; C25816; WMBP8F.
KW Late protein.
SQ SEQUENCE 280 AA; 29489 MW; 9B04A0B9BCDF4CF3 CRC64;

Query Match 96.4%; Score 27; DB 1; Length 280;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 96 GYAVEE 101

RESULT 15
VG85_BPPZA
ID VG85_BPPZA STANDARD; PRT; 280 AA.
AC P07532;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Head fiber protein (Late protein GP8.5).
GN Name=8.5;
OS Bacteriophage PZA.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14782; AAA32281.1; -.
DR PIR; C25816; WMBP8F.
KW Late protein.
SQ SEQUENCE 280 AA; 29489 MW; 9B04A0B9BCDF4CF3 CRC64;

Query Match 96.4%; Score 27; DB 1; Length 280;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 96 GYAVEE 101

RESULT 16
DAAA_STAHA
ID DAAA_STAHA STANDARD; PRT; 282 AA.
AC P54694;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE D-alanine aminotransferase (EC 2.6.1.21) (D-aspartate
DE aminotransferase) (D-amino acid aminotransferase) (D-amino acid
DE transaminase) (DAAT).
GN Name=Dat;
OS Staphylococcus haemolyticus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OC NCBI_TaxID=1283;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Y176;
RX MEDLINE=95113770; PubMed=7814322;
RA Pucci M.J., Thanassi J.A., Ho H.-T., Falk P.J., Dougherty T.J.;
RT "Staphylococcus haemolyticus contains two D-glutamic acid biosynthetic
RT activities, a glutamate racemase and a D-amino acid transaminase.";
RL J. Bacteriol. 177:336-342(1995).
CC -!- FUNCTION: Acts on the D-isomers of alanine, leucine, aspartate,
CC glutamate, aminobutyrate, norvaline and asparagine (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: D-alanine + 2-oxoglutarate = pyruvate + D-
CC glutamate.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- SIMILARITY: Belongs to the class-IV pyridoxal-phosphate-dependent
CC aminotransferase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U12238; AAA20396.1; -.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC phi-29-like viruses.
OC NCBI_TaxID=10757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87031573; PubMed=3095188;
RA Paces V., Vicek C., Urbanek P.;
RT "Nucleotide sequence of the late region of Bacillus subtilis phage
RT PZA, a close relative of phi 29.";
RL Gene 44:107-114(1986).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11813; AAA88485.1; -.
DR PIR; C24831; WMBP8H.
KW Late protein.
SQ SEQUENCE 280 AA; 29457 MW; 487C378FEE70339A CRC64;

Query Match 96.4%; Score 27; DB 1; Length 280;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 96 GYAVEE 101
```

```

DR HSSP; P19938; 1DAA.
DR InterPro; IPR001544; Aminotrans IV.
DR Pfam; PF01063; Aminotran 4; 1.
DR ProDom; PD001961; Aminotran 4; 1.
DR TIGRFAMs; TIGR01121; D.amino.aminoT; 1.
DR PROSITE; PS00770; AA_TRANSFERS_CLASS; 4; 1.
KW Amino transferase; Pyridoxal phosphate; Transferase.
FT BINDING 146 146 Pyridoxal phosphate (Potential).
SQ SEQUENCE 282 AA; 31769 MW; 43DEC87A570B4F13 CRC64;

Query Match 96.4%; Score 27; DB 1; Length 282;
Best Local Similarity 83.3%; Pred.No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 64 GYTVEE 69

RESULT 17
Q9HIB6 PRELIMINARY; PRT; 299 AA.
AC Q9HIB6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable transcription termination-antitermination factor related protein.
GN OrderedLocusNames=Tal425;
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum."
RL Nature 407:508-513 (2000).
DR EMBL; AL445067; CAC12545.1; -.
DR PIR; S11603; S11603.
DR GO; GO:0003711; F:transcriptional elongation regulator activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR005824; KOW.
DR InterPro; IPR006645; NGN.
DR InterPro; IPR011590; Ribosomal L26e.A.
DR InterPro; IPR003257; Ribosomal_NusG-.
DR InterPro; IPR008991; Transl_SH3-like.
DR Pfam; PF00467; KOW; 1.
DR ProDom; PD005267; Ribosomal_NusG; 1.
DR SMART; SM00739; KOW; 1.
DR SMART; SM00738; NGN; 1.
DR TIGRFAMs; TIGR00405; L26e_arch; 1.
KW Complete proteome.
SQ SEQUENCE 299 AA; 33396 MW; 4E2B3F86A3A26B52 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 299;
Best Local Similarity 83.3%; Pred.No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 58 GYSVEE 63

RESULT 18
Q9L973 PRELIMINARY; PRT; 306 AA.
ID Q9L973
AC Q9L973;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Lactococcus lactis.
OC Plasmid pCI2001.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=275;
RA Kearney K., Fitzgerald G.F.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179847; AAF27561.1; -.
DR InterPro; IPR005094; Relaxase.
DR Pfam; PF03432; Relaxase; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 306 AA; 36147 MW; 69E5182AB82C07A9 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 306;
Best Local Similarity 83.3%; Pred.No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 66 GYSVEE 71

RESULT 19
GPDA AQUAE STANDARD; PRT; 324 AA.
ID GPDA AQUAE
AC O67555;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94) (NAD(P)H-dependent glycerol-3-phosphate dehydrogenase).
GN Name=gpsA; OrderedLocusNames=AQ_1634;
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320; DOI=10.1038/32831;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olsen G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex aeolicus."
RL Nature 392:353-358 (1998).
DR -1- CATALYTIC ACTIVITY: Sn-glycerol 3-phosphate + NAD(P) (+) =
CC Glycerone phosphate + NAD(P)H.
DR -1- PATHWAY: De novo phospholipid biosynthesis; glycerol-3 phosphate
CC formation.
DR -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the NAD-dependent glycerol-3-phosphate
CC dehydrogenase family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000748; AAC07511.1; ALT_INIT.
CC HSSP; P90551; LEVI.
CC HMAP; MF_00394; -.
CC InterPro; IPR008927; 6GDGH_C_like.
CC InterPro; IPR006109; NAD_Gly3P_C.

```

```

DR InterPro; IPR006168; NAD_Gly3P_dh.
DR InterPro; IPR011128; NAD_Gly3P_dh_N.
DR Pfam; PF07479; NAD_Gly3P_dh_C; 1.
DR Pfam; PF01210; NAD_Gly3P_dh_N; 1.
DR PRINTS; PR00077; GPDHGRNASE.
DR ProDom; PD001278; NAD_Gly3P_dom; 1.
DR PROSITE; PS00957; NAD_G3PDH; 1.
DR Complete proteome; NAD; Oxidoreductase; Phospholipid biosynthesis.
SQ SEQUENCE 324 AA; 36032 MW; D2DBC4C7182CC500 CRC64;

Query Match 96.4%; Score 27; DB 1; Length 324;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 257 GYSVEE 262

RESULT 20
Q6FAC8 PRELIMINARY; PRT; 326 AA.
AC Q6FAC8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=ACIAD2186;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP SEQUENCE FROM N.A.
RA Barbe V., Vallent D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Ornston L.N., Weissbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 0:0-0(2004).
DR EMBL; CR543861; GAG68985.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 326 AA; 37784 MW; 7D00A904D0599DD4 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 326;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 254 GYSVEE 259

RESULT 21
Q8PSH1 PRELIMINARY; PRT; 327 AA.
AC Q8PSH1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein MW3108.
GN OrderedLocNames=MM3108;
OS Methanosarcina maei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goe1 / Goe1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierzer A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,

```

```

RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina maei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013568; AM32804.1; -.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. .; IEA.
DR InterPro; IPR003018; GAF.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR000700; PAS-assoc_C.
DR Pfam; PF00785; PAC; 1.
DR SMART; SM00065; GAF; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
DR PROSITE; PS01113; PAC; 1.
DR PROSITE; PS01112; PAS; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 327 AA; 37518 MW; 78B35D489BE7BE4D CRC64;

Query Match 96.4%; Score 27; DB 2; Length 327;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 68 GYTVEE 73

RESULT 22
Q8TMU0 PRELIMINARY; PRT; 330 AA.
AC Q8TMU0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein MA2561.
GN OrderedLocNames=MA2561;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA FitzHugh W., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.P., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010951; AA005943.1; -.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. .; IEA.
DR InterPro; IPR003018; GAF.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR000700; PAS-assoc_C.
DR Pfam; PF00785; PAC; 1.
DR SMART; SM00065; GAF; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
DR PROSITE; PS01113; PAC; 1.

```

```

DR PROSITE; PSS0112; PAS; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 330 AA; 37728 MW; 9E74D14625DC99C1 CRC64;

Query Match          96.4%; Score 27; DB 2; Length 330;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 71 GYTVEE 76

RESULT 23
Q98HF4 PRELIMINARY; PRT; 343 AA.
ID Q98HF4
AC Q98HF4
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DE Acetylpolylamine aminohydrolase.
GN OrderedLocusNames=mlr2893;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
CC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN SEQUENCE FROM N.A.
RC STRAIN=MAFF30309;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawaishima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
[1]
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=MAFF30309;
RX MEDLINE=21082936; PubMed=11214974;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawaishima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti (supplement).";
RL DNA Res. 7:381-406(2000).
DR EMBL; AP003000; BAB49912.1; -.
DR InterPro; IPR000286; Hlg_deacetylase.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
KW Complete proteome.
SQ SEQUENCE 343 AA; 37222 MW; FF962BAE96BCA318 CRC64;

Query Match          96.4%; Score 27; DB 2; Length 343;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 323 GYAVEE 328

RESULT 24
SYFA BACSU
ID SYFA BACSU STANDARD; PRT; 344 AA.
AC P17921; P94539;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 01-OCT-2004 (Rel. 45; Last annotation update)

```

```

DE Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
DE (Phenylalanine--tRNA ligase alpha chain) (PheRS).
GN Name=pheS; OrderedLocusNames=BSU28640;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=91175935; PubMed=2127701;
RA Brakhage A., Wozny M., Putzer H.;
RT "Structure and nucleotide sequence of the Bacillus subtilis
RT phenylalanyl-tRNA synthetase genes.";
RL Biochimie 72:725-734(1990).
[2]
RN ERRATUM.
RX MEDLINE=91234765; PubMed=1903307;
RA Brakhage A., Wozny M., Putzer H.;
RL Biochimie 73:127-127(1991).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124191; PubMed=8969504;
RA Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K., Sanders J.,
RA Emerson P.T., Harwood C.R.;
RT "The dnaB-pheA (256 degrees-240 degrees) region of the Bacillus
RT subtilis chromosome containing genes responsible for stress responses,
RT the utilization of plant cell walls and primary metabolism.";
RL Microbiology 142:3067-3078(1996).
[4]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conner I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dunthof A., Ehrlich S.D., Emerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Fertari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Coffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC Phe-tRNA synthetase alpha chain subfamily 1.

```



```

RESULT 27
Q88C43 ID Q88C43 PRELIMINARY; PRT; 347 AA.
AC Q88C43;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Acetylpolylamine aminohydrolase.
GN OrderedLocusNames=PP5340;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzez A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoef A., Tuemmli B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AF016794; AAN70905.1; -.
DR TIGR; PP5340; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000286; His deacetylase.
DR Pfam; PF00850; Hist deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 347 AA; 37880 MW; 39DAE40CA46D0877 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 347;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 322 GYAVEE 327

RESULT 28
GBT_XENLA ID GBT_XENLA STANDARD; PRT; 349 AA.
AC P38407;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Guanine nucleotide-binding protein G(t), alpha subunit (Transducin
DE alpha chain).
GN Name=GNAT;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Knox B.E., Scalzetti L.C., Batni S., Wang J.Q.;
RT "Molecular cloning of the abundant rhodopsin and transducin from
RT Xenopus laevis.";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are
CC involved as modulators or transducers in various transmembrane
CC signaling systems. Transducin is an amplifier and one of the
CC transducers of a visual impulse that performs the coupling between

```

```

rhodopsin and cGMP-phosphodiesterase.
!- SUBUNIT: G proteins are composed of 3 units; alpha, beta and
gamma. The alpha chain contains the guanine nucleotide binding
site.
!- SIMILARITY: Belongs to the G-alpha family. Subfamily 1
(G(i/o/t/z)).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; L07771; AAA88693.1; -.
DR HSSP; P04695; 1TND.
DR InterPro; IPR001019; Gprotein_alpha.
DR InterPro; IPR001408; Gprotein_alphai.
DR InterPro; IPR011025; Transducn_inser.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR PRINTS; PR00441; GPROTEINAI.
DR ProDom; PD000281; Gprotein_alpha; 1.
KW ADP-ribosylation; GTP-binding; Lipoprotein; Myristate; Transducer;
KW Vision.
FT INIT_MBT 0 0 By similarity.
FT LIPID 1 1 N-myristoyl glycine (By similarity).
FT NP_BIND 35 42 GTP (By similarity).
FT NP_BIND 195 199 GTP (By similarity).
FT NP_BIND 264 267 GTP (By similarity).
FT MOD_RES 173 173 ADP-ribosylarginine (by cholera toxin)
FT MOD_RES 346 346 (By similarity).
FT MOD_RES 346 346 ADP-ribosylcysteine (by pertussis toxin)
FT MOD_RES 346 346 (By similarity).
SQ SEQUENCE 349 AA; 39679 MW; 76CBF22E81EDA7D4 CRC64;

Query Match 96.4%; Score 27; DB 1; Length 349;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 55 GYSVEE 60

RESULT 29
Q6MX97 ID Q6MX97 PRELIMINARY; PRT; 350 AA.
AC Q6MX97;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN ORFNames=SMR0304;
OS Serratia marcescens.
OG Plasmid R478.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RA Gilmour M.W., Thomson N.R., Saunders M., Parkhill J., Taylor D.E.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX664015; CAES1827.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 350 AA; 39346 MW; 2BB4A69300A78352 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 350;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

```

```
Db 135 GYSVEE 140
|| |||
RESULT 30
CAE51827 PRELIMINARY; PRT; 350 AA.
AC CAE51827;
DT 02-MAR-2004 (TRENBLrel. 27, Created)
DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN SMR0304.
OS Serratia marcescens.
OG Flaemid R478.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RA Gilmore M.W., Thomson N.R., Saunders M., Parkhill J., Taylor D.E.;
RT "The complete nucleotide sequence of the resistance plasmid
RT R478: defining the backbone components of incompatibility group H
RT conjugative plasmids through comparative genomics.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BX664015; CAE51827.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 350 AA; 39346 MW; 2B54A69300A78352 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 350;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
|| |||
Db 135 GYSVEE 140

RESULT 31
Q6MPJ7 PRELIMINARY; PRT; 354 AA.
AC Q6MPJ7;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Carbamoyl-phosphate synthase small chain (EC 6.3.5.5).
GN Name=carA; OrderedLocNames=BQ0856;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RA Rendluc S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goesmann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
DR EMBL: BX842648; CAE78801.1; -.
KW Ligase.
SQ SEQUENCE 354 AA; 39255 MW; A1E0C9DF8C9ED85A CRC64;

Query Match 96.4%; Score 27; DB 2; Length 354;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
|| |||
Db 292 GYAVEE 297

RESULT 32
CAE78801 PRELIMINARY; PRT; 354 AA.
ID CAE78801;
AC CAE78801;
DT 02-MAR-2004 (TRENBLrel. 27, Created)
DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)
DE Carbamoyl-phosphate synthase small chain (EC 6.3.5.5).
GN CARA OR BD0856.
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164;
RA Rendluc S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goesmann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
DR EMBL: BX842648; CAE78801.1; -.
KW Ligase.
SQ SEQUENCE 354 AA; 39255 MW; A1E0C9DF8C9ED85A CRC64;

Query Match 96.4%; Score 27; DB 2; Length 354;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
|| |||
Db 292 GYAVEE 297

RESULT 33
CARA_THETN STANDARD; PRT; 356 AA.
ID CARA_THETN;
AC Q8REK1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Carbamoyl-phosphate synthase small chain (EC 6.3.5.5) (Carbamoyl-
DE phosphate synthetase glutamine chain).
GN Name=carA; OrderedLocNames=TFE0815;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
```



Genome Res. 12:689-700 (2002).  
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +  
CC phosphate + L-glutamate + carbamoyl phosphate.  
CC -!- PATHWAY: Arginine biosynthesis.  
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.  
CC -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain  
CC promotes the hydrolysis of glutamine to ammonia, which is used by  
CC the large (or ammonia) chain to synthesize carbamoyl phosphate (by  
CC similarity).  
CC -!- SIMILARITY: Belongs to the carA family.  
CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE013048; AAM24072.1; -.  
DR HSSP; P00907; 1CE8.  
DR HAMAP; MF\_01209; -. 1.  
DR InterPro; IPR006274; CarA\_synth\_small.  
DR InterPro; IPR001317; CP\_synthGATase.  
DR InterPro; IPR002474; CP\_synthsmall.  
DR InterPro; IPR000991; GATase\_1.  
DR Pfam; PF00988; CPsase\_sm\_chain; 1.  
DR Pfam; PF00117; GATase; 1.  
DR PRINTS; PR00059; CPSCATASE.  
DR PRINTS; PR00096; GATASE.  
DR TIGRFAMs; TIGR01368; CPsaseIleSmall; 1.  
DR PROSITE; PS00442; GATASE\_TYPE\_I; 1.  
KW Arginine biosynthesis; Complete proteome; Glutamine amidotransferase;  
KW Ligase; Pyrimidine biosynthesis.  
FT DOMAIN 1 159 CPsase.  
FT DOMAIN 160 356 Glutamine amidotransferase.  
FT ACT\_SITE 238 238 GATase (by similarity).  
SQ SEQUENCE 356 AA; 39975 MW; 66B5BB8C91161CCB CRC64;  
  
Query Match 96.4%; Score 27; DB 1; Length 356;  
Best Local Similarity 83.3%; Pred. No. 4.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVEE 6  
DB 282 GYAVEE 287  
[1] |||||  
[3] |||||  
SEQUENCE FROM N.A.  
Pain A., Woodward J., Quail M.A., Anderson M.J., Clark R., Collins M.,  
Fosker N., Fraser A., Harris D., Larke N., Murphy L., Humphray S.,  
O'Neil S., Perteau M., Price C., Rabinovitch E., Rajandream M.A.,  
Salzberg S., Saunders D., Seegar K., Sharp S., Warren T.,  
Denning D.W., Barrell B., Hall N.;  
"Insight into the genome of Aspergillus fumigatus: analysis of a 922  
kb region encompassing the nitrate assimilation gene cluster.";  
Fungal Genet. Biol. 41:443-453(2004).  
RL EMBL; AL713629; CAD28435.1; -.  
DR EMBL; AJ277436; CAC81805.1; -.  
DR EMBL; BX649606; CAF32008.1; -.  
DR HSSP; P10824; 1BOF.  
DR GO; GO:0005525; F:GTP binding; IEA.  
DR GO; GO:0004871; F:signal transducer activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR002975; Fungi\_Gproteina.  
DR InterPro; IPR002975; Fungi\_Gproteina.  
DR InterPro; IPR001019; Gproteina\_alpha.  
DR InterPro; IPR011025; Transducn\_insert.  
DR Pfam; PF00503; G-alpha; 1.  
DR PRINTS; PR00318; GPROTEINA.  
DR PRINTS; PR01241; GPROTEINAFNG.  
DR ProDom; PD000281; Gproteina\_alpha; 1.  
DR SMART; SM00275; G\_alpha; 1.  
SQ SEQUENCE 356 AA; 41039 MW; 8DF1F9BE4D148CB6 CRC64;

DR InterPro; IPR001019; Gproteina\_alpha.  
DR InterPro; IPR011025; Transducn\_insert.  
DR Pfam; PF00503; G-alpha; 1.  
DR PRINTS; PR00318; GPROTEINA.  
DR PRINTS; PR01241; GPROTEINAFNG.  
DR ProDom; PD000281; Gproteina\_alpha; 1.  
DR SMART; SM00275; G\_alpha; 1.  
SQ SEQUENCE 356 AA; 40958 MW; B6B4C473852B3A51 CRC64;  
  
Query Match 96.4%; Score 27; DB 2; Length 356;  
Best Local Similarity 83.3%; Pred. No. 4.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVEE 6  
DB 63 GYTVEE 68  
[1] |||||  
[2] |||||  
[3] |||||  
SEQUENCE FROM N.A.  
Knowles D.G., Warren T., Hall N., Quail M., Woodward J.R.,  
Denning D.W., Anderson M.J., Barrell B.;  
Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
[2]  
[2]  
SEQUENCE FROM N.A.  
RC STRAIN=ATCC46645;  
RA Liebmam B., Gattung S., Brakhage A.A.;  
"cAMP signaling in Aspergillus fumigatus is involved in the regulation  
of the virulence gene pksp and in defense against killing by  
macrophages.";  
Mol. Genet. Genomics 269:420-435(2003).  
RN [3]  
SEQUENCE FROM N.A.  
RX PubMed=14998527;  
RA Pain A., Woodward J., Quail M.A., Anderson M.J., Clark R., Collins M.,  
Fosker N., Fraser A., Harris D., Larke N., Murphy L., Humphray S.,  
O'Neil S., Perteau M., Price C., Rabinovitch E., Rajandream M.A.,  
Salzberg S., Saunders D., Seegar K., Sharp S., Warren T.,  
Denning D.W., Barrell B., Hall N.;  
"Insight into the genome of Aspergillus fumigatus: analysis of a 922  
kb region encompassing the nitrate assimilation gene cluster.";  
Fungal Genet. Biol. 41:443-453(2004).  
RL EMBL; AL713629; CAD28435.1; -.  
DR EMBL; AJ277436; CAC81805.1; -.  
DR EMBL; BX649606; CAF32008.1; -.  
DR HSSP; P10824; 1BOF.  
DR GO; GO:0005525; F:GTP binding; IEA.  
DR GO; GO:0004871; F:signal transducer activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR002975; Fungi\_Gproteina.  
DR InterPro; IPR002975; Fungi\_Gproteina.  
DR InterPro; IPR001019; Gproteina\_alpha.  
DR InterPro; IPR011025; Transducn\_insert.  
DR Pfam; PF00503; G-alpha; 1.  
DR PRINTS; PR00318; GPROTEINA.  
DR PRINTS; PR01241; GPROTEINAFNG.  
DR ProDom; PD000281; Gproteina\_alpha; 1.  
DR SMART; SM00275; G\_alpha; 1.  
SQ SEQUENCE 356 AA; 41039 MW; 8DF1F9BE4D148CB6 CRC64;

```
Query Match          96.4%; Score 27; DB 2; Length 356;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 63 GYTVEE 68

RESULT 36
CAF32008 PRELIMINARY; PRT; 356 AA.
AC CAF32008;
DT 13-MAY-2004 (TrEMBLrel. 27, Created)
DT 13-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 13-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE G-protein alpha subunit, putative.
GN GANB OR AFA5C11.10C.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;
RN [1]
RP SEQUENCE FROM N.A.
RA Pain A., Woodward J., Quail M.A., Anderson M.J., Clark R., Collins M.,
RA Foeker N., Fraser A., Harris D., Larke N., Murphy L., Humphray S.,
RA O'Neil S., Perteaux M., Price C., Rabinowitch E., Rajandream M.A.,
RA Salzberg S., Saunders D., Seegar K., Sharp S., Warren T.,
RA Denning D.W., Barrell B., Hall N.;
RA "Insight into the genome of Aspergillus fumigatus: analysis of a 922
RT kb region encompassing the nitrate assimilation gene cluster.";
RL Fungal Genet. Biol. 41:443-453(2004).
DR EMBL; BX649606; CAF32008.1; -.
SQ SEQUENCE 356 AA; 41039 MW; 8DF1F8BEAD148CB6 CRC64;

Query Match          96.4%; Score 27; DB 2; Length 356;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 63 GYTVEE 68

RESULT 37
Q729K5 PRELIMINARY; PRT; 358 AA.
AC Q729K5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE G protein alpha subunit.
GN Name=gasc;
OS Penicillium marneffei.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=37727;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22692143; PubMed=12807770;
RA Zuber S., Hynes M.J., Adrianopoulos A.;
RT "The G-protein alpha-subunit Gasc plays a major role in germination in
RT the dimorphic fungus Penicillium marneffei.";
RL Genetics 164:487-499(2003).
DR EMBL; AY170625; AAO41857.1; -.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha; 1.
DR ProDom; PD000281; Gprotein_alpha; 1.
DR SMART; SM00275; G_alpha; 1.
SQ SEQUENCE 358 AA; 41332 MW; 17B1C358D5B06ADC CRC64;
```

```
Query Match          96.4%; Score 27; DB 2; Length 358;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 63 GYSVEE 68

RESULT 38
Q9YD80 PRELIMINARY; PRT; 363 AA.
AC Q9YD80;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE1032.
GN OrderedLocusNames=APE1032;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jinno K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA Hotoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA80017.1; -.
PIR; A72702; A72702.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 363 AA; 39437 MW; 674F2FDD611901B7 CRC64;

Query Match          96.4%; Score 27; DB 2; Length 363;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 6 GYAVEE 11

RESULT 39
Q6HES7 PRELIMINARY; PRT; 365 AA.
AC Q6HES7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Carbanoyl-phosphate synthase, small subunit (EC 6.3.5.5).
GN Name=carA, ORFNames=Bt9727_3629;
OS Bacillus thuringiensis serovar konkukian str. 97-27.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus thuringiensis serovar konkukian.
OX NCBI_TaxID=281309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Brettn T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017355; AAT61300.1; -.
DR InterPro; IPR006220; Anth_synthII.
DR InterPro; IPR006274; CarA_synth_small.
```

DR InterPro; IPR001317; CP\_synthGATase.  
 DR InterPro; IPR002474; CP\_synthsmall.  
 DR InterPro; IPR000991; GATase\_1.  
 DR Pfam; PF00988; CPSase\_sm\_chain; 1.  
 DR Pfam; PF00117; GATase; 1.  
 DR PRINTS; PR00097; ANTSNTHASEII.  
 DR PRINTS; PR00099; CPSGATASE.  
 DR PRINTS; PR00096; GATASE.  
 DR TIGRFAMs; TIGR01368; CPSaseIIsmall; 1.  
 DR PROSITE; PS00442; GATASE\_TYPE\_I; 1.  
 KW Ligase.  
 SQ SEQUENCE 365 AA; 40376 MW; 0F340C1663AF29EF CRC64;

Query Match 96.4%; Score 27; DB 2; Length 365;  
 Best Local Similarity 83.3%; Pred. No. 4.6e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 |||||

Db 289 GYAVEE 294

RESULT 40

ID Q73212 PRELIMINARY; PRT; 365 AA.

AC Q73212;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Carbamoyl-phosphate synthase, small subunit (EC 6.3.5.5).

GN Name=carA; OrderedLocusNames=BCE3932;  
 OS Bacillus cereus (strain ATCC 10987);  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=222523;

FN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=14960714;

RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,  
 RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,  
 RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;  
 RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic  
 RT adaptations and a large plasmid related to Bacillus anthracis pX01";  
 RL Nucleic Acids Res. 32:977-988(2004).

DR EMBL; AE017276; AAS42835.1; -;  
 DR TIGR; BCE3932; -;  
 DR GO; GO:0004089; F:carbamoyl-phosphate synthase (glutamine-hyd. . .; IEA.  
 DR GO; GO:0016874; F:ligase activity; IEA.

DR InterPro; IPR006220; Anch\_synthII.  
 DR InterPro; IPR006274; CarA\_synth\_small.  
 DR InterPro; IPR001317; CP\_synthGATase.  
 DR InterPro; IPR002474; CP\_synthsmall.  
 DR InterPro; IPR000991; GATase\_1.  
 DR Pfam; PF00988; CPSase\_sm\_chain; 1.  
 DR Pfam; PF00117; GATase; 1.  
 DR PRINTS; PR00097; ANTSNTHASEII.  
 DR PRINTS; PR00099; CPSGATASE.  
 DR PRINTS; PR00096; GATASE.  
 DR TIGRFAMs; TIGR01368; CPSaseIIsmall; 1.  
 DR PROSITE; PS00442; GATASE\_TYPE\_I; 1.  
 KW Complete proteome; Ligase.

SQ SEQUENCE 365 AA; 40376 MW; 0F31591333FA2CBA CRC64;

Query Match 96.4%; Score 27; DB 2; Length 365;  
 Best Local Similarity 83.3%; Pred. No. 4.6e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
 |||||

Db 289 GYAVEE 294

RESULT 41

Q81WF1

ID Q81WF1 PRELIMINARY; PRT; 365 AA.  
 AC Q81WF1; QGHUJ9; Q6KNT4;  
 DT 01-JUN-2003 (TReMBLrel. 24, Created)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
 DT 01-OCT-2004 (TReMBLrel. 28, Last annotation update)  
 DE Carbamoyl-phosphate synthase, small subunit.  
 GN Name=carA; OrderedLocusNames=BA4026, BAS3738; ORFNames=GBAA4026;  
 OS Bacillus anthracis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=11392;  
 FN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Ames / isolate Porton;  
 RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;  
 RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,  
 RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,  
 RA Holtzapple E.K., Oekstad O.A., Helgason E., Rilstone J., Wu M.,  
 RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,  
 RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,  
 RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,  
 RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,  
 RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,  
 RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,  
 RA Salzberg S.L., Thomason B., Friedlander A.M., Koehler T.M.,  
 RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;  
 RT "The genome sequence of Bacillus anthracis Ames and comparison to  
 RT closely related bacteria";  
 RL Nature 423:81-86(2003).  
 FN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Ames / isolate 0581;  
 RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,  
 RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,  
 RA Fraser C.M.;  
 RT "Bacillus anthracis comparative genomics";  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 FN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sterner;  
 RA Bretton T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,  
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,  
 RA Richardson P., Rubin E., Rice H.,  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE017036; AAP27753.1; -;  
 DR EMBL; AE017334; AAT33143.1; -;  
 DR EMBL; AE017225; AAT56040.1; -;  
 DR HSP; P00907; 1CE8.  
 DR TIGR; BA4026; -;  
 DR GO; GO:0004049; F:anthranilate synthase activity; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004086; F:carbamoyl-phosphate synthase activity; IEA.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR GO; GO:0006807; P:nitrogen metabolism; IEA.  
 DR InterPro; IPR006220; Anch\_synthII.  
 DR InterPro; IPR006274; CarA\_synth\_small.  
 DR InterPro; IPR001317; CP\_synthGATase.  
 DR InterPro; IPR002474; CP\_synthsmall.  
 DR InterPro; IPR000991; GATase\_1.  
 DR Pfam; PF00988; CPSase\_sm\_chain; 1.  
 DR Pfam; PF00117; GATase; 1.  
 DR PRINTS; PR00097; ANTSNTHASEII.  
 DR PRINTS; PR00099; CPSGATASE.  
 DR PRINTS; PR00096; GATASE.  
 DR TIGRFAMs; TIGR01368; CPSaseIIsmall; 1.  
 DR PROSITE; PS00442; GATASE\_TYPE\_I; 1.  
 SQ SEQUENCE 365 AA; 40376 MW; 0F340C1663AF29EF CRC64;

Query Match 96.4%; Score 27; DB 2; Length 365;  
 Best Local Similarity 83.3%; Pred. No. 4.6e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6

Db	289	GYAVEE 294		
RESULT 42				
AA342835				
ID	AA342835	PRELIMINARY;	PRT;	365 AA.
AC	AA342835;			
DT	02-MAR-2004	(TrEMBLrel. 27, Created)		
DT	02-MAR-2004	(TrEMBLrel. 27, Last sequence update)		
DT	04-MAY-2004	(TrEMBLrel. 27, Last annotation update)		
DE	Carbamoyl-phosphate synthase, small subunit (EC 6.3.5.5).			
GN	CARA OR BCE3932.			
OS	Bacillus cereus (strain ATCC 10987).			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=222523;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=14960714;			
RA	Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L., Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F., Nelson W.C., Kolstoe A.B., Fraser C.M., Read T.D.,			
RA	"The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic adaptations and a large plasmid related to Bacillus anthracis pXOI."			
RT	Nucleic Acids Res. 32:977-988(2004).			
RL	EMBL; AE017276; AA342835.1; -.			
DR	TIGR; BCE3932; -.			
KW	Ligase.			
SQ	SEQUENCE 365 AA; 40376 MW; 0F31591333FA2CBA CRC64;			
Query Match	96.4%;	Score 27;	DB 2;	Length 365;
Best Local Similarity	83.3%;	Pred. No. 4.6e+02;		
Matches	5; Conservative	0; Mismatches	1; Indels	0; Gaps
QY	1 GYXVEE 6			
Db	289	GYAVEE 294		
RESULT 43				
AAT33143				
ID	AAT33143	PRELIMINARY;	PRT;	365 AA.
AC	AAT33143;			
DT	01-JUN-2004	(TrEMBLrel. 27, Created)		
DT	01-JUN-2004	(TrEMBLrel. 27, Last sequence update)		
DT	01-JUN-2004	(TrEMBLrel. 27, Last annotation update)		
DE	Carbamoyl-phosphate synthase, small subunit.			
GN	CARA OR GBAA4026.			
OS	Bacillus anthracis str. Ames 0581.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;			
OC	Bacillus cereus group; Bacillus anthracis.			
OX	NCBI_TaxID=261594;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	STRAIN=Ames 0581;			
RA	Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B., Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;			
RT	"Bacillus anthracis comparative Genomics."			
RL	Submitted (MAY-2004) to the EMBL/GenBank/DDJ databases.			
DR	EMBL; AE017334; AAT33143.1; -.			
SQ	SEQUENCE 365 AA; 40376 MW; 0F340C1663AF29EF CRC64;			
Query Match	96.4%;	Score 27;	DB 2;	Length 365;
Best Local Similarity	83.3%;	Pred. No. 4.6e+02;		
Matches	5; Conservative	0; Mismatches	1; Indels	0; Gaps
QY	1 GYXVEE 6			
Db	289	GYAVEE 294		
RESULT 44				
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Created)		
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			
AC	Q81982;			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
SERC_METAC				
ID	Q81982	PRELIMINARY;	PRT;	373 AA.
ID	Q81982			

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Carbamoyl-phosphate synthase small chain (EC 6.3.5.5).  
 GN ORFNames=BC3887;  
 OS Bacillus cereus (strain ATCC 14579 / DSM 31).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 ON NCBI\_TaxID=226900;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;  
 RX Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,  
 RA Kapatral V., Bhattacharya A., Resnik G., Mikhailova N., Lapidus A.,  
 RA Chu L., Mazur M., Goltsev E., Larsen N., D'Souza M., Walunas T.,  
 RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,  
 RA Overbeek R., Kyrpides N.C.;  
 RT "Genome sequence of *Bacillus cereus* and comparative analysis with  
 RT *Bacillus anthracis*.";  
 RL Nature 423:87-91(2003).  
 DR EMBL; AE017010; AAP10808.1; --  
 DR HSSP; P00907; 1CS8.  
 DR GO; GO:0004049; F:anthranilate synthase activity; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004088; F:carbamoyl-phosphate synthase (glutamine-hyd. . .); IEA.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0009059; P:biosynthesis; IEA.  
 DR GO; GO:0006807; P:nitrogen metabolism; IEA.  
 DR InterPro; IPR006220; Antc\_synthIII.  
 DR InterPro; IPR006274; CarA\_synth\_small.  
 DR InterPro; IPR001317; CP\_synthGATase.  
 DR InterPro; IPR002474; CP\_synthGATase.  
 DR InterPro; IPR000991; GATase\_1.  
 DR Pfam; PF00988; CPSase\_sm\_chain; 1.  
 DR Pfam; PF00117; GARase; 1.  
 DR PRINTS; PR00097; ANTSNTNTHASE.  
 DR PRINTS; PR00099; CPSGATASE.  
 DR PRINTS; PR00096; GATASE.  
 DR TIGRFAMs; TIGR01368; CPSaseIIsmall; 1.  
 DR PROSITE; PS00442; GATASE\_TYPE\_I; 1.  
 KW Ligase.  
 SQ SEQUENCE 373 AA; 41181 MW; E22EBDA2C21724A6 CRC64;  
  
 Query Match 96.4%; Score 27; DB 2; Length 373;  
 Best Local Similarity 83.3%; Pred. No. 4.7e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 GYXVEE 6  
 Db 297 GYAVEE 302  
  
 RESULT 46  
 C1SY PYRFU STANDARD; PRT; 376 AA.  
 AC Q53554;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Citrate synthase (EC 2.3.3.1).  
 GN Name=gltA; OrderedLocusNames=PF0203;  
 OS Pyrococcus furiosus.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 ON NCBI\_TaxID=2261;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=96081442; PubMed=8532683;  
 RA Muir J.M., Russell R.J., Hough D.W., Danson M.J.;  
 RT "Citrate synthase from the hyperthermophilic Archaeon, *Pyrococcus*  
 RT *furiosus*.";  
 RL Protein Eng. 8:583-592(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;

"The complete sequence of the *Pyrococcus furiosus* genome.";  
 Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RT X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=97400454; PubMed=9254593;  
 RA Russell R.J., Ferguson J.M., Hough D.W., Danson M.J., Taylor G.L.;  
 RT "The crystal structure of citrate synthase from the hyperthermophilic  
 RT archaeon *Pyrococcus furiosus* at 1.9-A resolution.";  
 RL Biochemistry 36:9983-9994(1997).  
 CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate = citrate +  
 CC CoA.  
 CC -!- PATHWAY: Tricarboxylic acid cycle.  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SIMILARITY: Belongs to the citrate synthase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; S81109; AAB35835.2; --  
 DR EMBL; AE010146; AAL80327.1; --  
 DR PDB; 1AJ8; X-ray; A/B=--  
 DR InterPro; IPR002020; Citrate\_synth.  
 DR Pfam; PF00285; Citrate\_synth; 1.  
 DR PRINTS; PR00143; CITRNTNTHASE.  
 DR TIGRFAMs; TIGR01798; cit\_synth\_I; 1.  
 DR TIGRFAMs; TIGR01800; cit\_synth\_II; 1.  
 DR PROSITE; PS00480; CITRATE\_SYNTHASE; 1.  
 DR 3D-structure; Allosteric enzyme; Complete proteome;  
 KW Direct protein sequencing; Transferase; Tricarboxylic acid cycle.  
 FT INIT MET 0 0  
 FT ACT SITE 262 262 By similarity.  
 FT ACT\_SITE 312 312 By similarity.  
 FT HELIX 8 10  
 FT TURN 11 12  
 FT STRAND 14 18  
 FT STRAND 21 24  
 FT TURN 25 28  
 FT STRAND 29 32  
 FT TURN 33 34  
 FT STRAND 35 36  
 FT HELIX 37 43  
 FT HELIX 46 55  
 FT HELIX 61 72  
 FT TURN 73 74  
 FT HELIX 79 87  
 FT TURN 90 91  
 FT HELIX 94 108  
 FT TURN 110 113  
 FT HELIX 118 143  
 FT TURN 144 145  
 FT TURN 153 154  
 FT HELIX 157 166  
 FT HELIX 172 185  
 FT HELIX 192 201  
 FT TURN 202 204  
 FT HELIX 207 219  
 FT TURN 221 225  
 FT HELIX 226 237  
 FT HELIX 240 242  
 FT HELIX 243 253  
 FT TURN 254 254  
 FT TURN 258 259  
 FT STRAND 260 260  
 FT HELIX 270 282  
 FT HELIX 285 301  
 FT TURN 302 306  
 FT STRAND 309 309  
 FT TURN 311 314

FT HELIX 315 320  
 FT TURN 321 323  
 FT HELIX 326 328  
 FT HELIX 329 349  
 FT TURN 350 351  
 FT STRAND 358 361  
 FT STRAND 367 367  
 FT HELIX 372 374  
 SQ SEQUENCE 376 AA; 42918 MW; AD466B1F0002DD82 CRC64;

Query Match 96.4%; Score 27; DB 1; Length 376;  
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 34 GYSVEE 39

RESULT 47  
 Q8TZ62 PRELIMINARY; PRT; 385 AA.

AC Q8TZ62: (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Predicted GTPase or GTP-binding protein.  
 GN OrderedLocNames=MK0077;  
 OS Methanopyrus kandleri.  
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
 OC Methanopyrus.  
 OX NCBI\_TaxID=2320;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;  
 RA MEDLINE=21927647; PubMed=11930014;  
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,  
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
 RA Malykh A.G., Koonin E.V., Kozhavkin S.A.;  
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
 and monophyly of archaeal methanogens.";  
 RL EMBL; AE010308; AAM01294.1; -.  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
 DR GO; GO:0005786; C:signal recognition particle; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0006614; P:SRP-dependent cotranslational membrane targ. .; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR000897; SRP54.  
 DR ProDom; PD000819; SRP54; 1.  
 DR SMART; SM00382; AAA; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 385 AA; 41992 MW; F8D314B15034E321 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 385;  
 Best Local Similarity 83.3%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 10 GYSVEE 15

RESULT 48  
 Q94CF8 PRELIMINARY; PRT; 385 AA.

AC Q94CF8:  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Branched-chain amino acid aminotransferase.  
 OS Capsicum annuum (Bell pepper).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Capsicum.  
 OX NCBI\_TaxID=4072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Balducci E., Richins R.D., Curry J., O'Connell M.A.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY034379; AAK57535.1; -.  
 DR HSP; O15382; IKT8.  
 DR GO; GO:0004084; F:branched-chain-amino-acid transaminase acti. .; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0009081; P:branched chain family amino acid metabolism; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001544; Aminotrans\_IV.  
 DR InterPro; IPR005786; B\_amino\_transII.  
 DR Pfam; PF01063; Aminotran\_4; 1.  
 DR ProDom; PD001961; Aminotrans\_IV; 1.  
 DR TIGRFAMs; TIGR01123; lIve II; 1.  
 DR PROSITE; PS00770; AA\_TRANSFER\_CLASS\_4; 1.  
 KW Aminotransferase; Transferase.  
 SQ SEQUENCE 385 AA; 42459 MW; 8F74581D8D0C201D CRC64;

Query Match 96.4%; Score 27; DB 2; Length 385;  
 Best Local Similarity 83.3%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 |||||  
 Db 301 GYTVEE 306

RESULT 49

ID Q20910 PRELIMINARY; PRT; 390 AA.  
 AC Q20910:  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein F56H9.4.  
 GN Name=F56H9.4;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 investigating biology.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Burton J.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z74473; CAA98953.1; -.  
 DR PIR; T22810; T22810.  
 DR HSP; P04896; 1CUL.  
 DR WormPep; F56H9.4; CE20886.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
 DR InterPro; IPR001019; Gprotein\_alpha.  
 DR InterPro; IPR001025; Transducn\_insert.  
 DR Pfam; PF00503; G-alpha; 1.  
 DR PRINTS; PR00318; GPROTEINA.  
 DR SMART; SM00275; G\_alpha; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 390 AA; 45205 MW; F8773E2A01FAB093 CRC64;

```

Query Match          96.4%; Score 27; DB 2; Length 390;
Best Local Similarity 83.3%; Pred.No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Oy 1 GYXVEE 6
Db 71 GYSVEE 76

RESULT 50
O72TP8 PRELIMINARY; PRT; 390 AA.
AC O72TP8;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Alpha-methylacyl-CoA racemase.
GN OrderedLocusNames=L1C10969;
OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar
OS Copenhageni).
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.
OX NCBI_TaxID=44275;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fiocruz Li-130;
RX PubMed=15028702;
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,
RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,
RA Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorri H.,
RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Giglioti E.A.,
RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
RA Jeronimo S.M.B., Junqueira-de-Azevedo I.L.M., Kimura E.T.,
RA Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
RA de Oliveira R.C., Pereira G.G., Reis M.S., Schrieffer A.,
RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
RT "Comparative genomics of two Leptospira interrogans serovars reveals
RT novel insights into physiology and pathogenesis.";
RL J. Bacteriol. 186:2164-2172(2004).
DR EMBL; AE017290; AAS69580.1; -
DR InterPro; IPR003673; CAIB-BAIF.
DR Pfam; PF02515; CoA_transf_3; 1.
KW Complete proteome.
SQ SEQUENCE 390 AA; 43625 MW; 7F1C916C80423807 CRC64;

Query Match          96.4%; Score 27; DB 2; Length 390;
Best Local Similarity 83.3%; Pred.No. 5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVEE 6
Db 375 GYTVEE 380

RESULT 51
O8F1J1 PRELIMINARY; PRT; 390 AA.
AC O8F1J1;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Alpha-methylacyl-CoA racemase (EC 5.1.99.4).
GN OrderedLocusNames=L31144;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,

```

```

RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-O., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AE011477; AAN50342.1; -
DR GO; GO:0008111; F:alpha-methylacyl-CoA racemase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CAIB-BAIF.
DR Pfam; PF02515; CoA_transf_3; 1.
KW Complete proteome; Isomerase.
SQ SEQUENCE 390 AA; 43625 MW; 7F1C916C80423807 CRC64;

Query Match          96.4%; Score 27; DB 2; Length 390;
Best Local Similarity 83.3%; Pred.No. 5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVEE 6
Db 375 GYTVEE 380

RESULT 52
AAS69580 PRELIMINARY; PRT; 390 AA.
ID AAS69580;
AC AAS69580;
DT 24-MAR-2004 (TRENBLrel. 27, Created)
DT 24-MAR-2004 (TRENBLrel. 27, Last sequence update)
DT 11-MAY-2004 (TRENBLrel. 27, Last annotation update)
DE Alpha-methylacyl-CoA racemase.
GN L1C10969;
OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar
OS Copenhageni).
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.
OX NCBI_TaxID=44275;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fiocruz Li-130;
RX PubMed=15028702;
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,
RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,
RA Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorri H.,
RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Giglioti E.A.,
RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
RA Jeronimo S.M.B., Junqueira-de-Azevedo I.L.M., Kimura E.T.,
RA Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
RA de Oliveira R.C., Pereira G.G., Reis M.S., Schrieffer A.,
RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
RT "Comparative genomics of two Leptospira interrogans serovars reveals
RT novel insights into physiology and pathogenesis.";
RL J. Bacteriol. 186:2164-2172(2004).
DR EMBL; AE017290; AAS69580.1; -
DR InterPro; IPR003673; CAIB-BAIF.
DR Pfam; PF02515; CoA_transf_3; 1.
KW Complete proteome.
SQ SEQUENCE 390 AA; 43625 MW; 7F1C916C80423807 CRC64;

Query Match          96.4%; Score 27; DB 2; Length 390;
Best Local Similarity 83.3%; Pred.No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GYXVEE 6
Db 375 GYTVEE 380

RESULT 53
Q98PFI PRELIMINARY; PRT; 398 AA.
ID Q98PFI

```

AC Q98PF1;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE NADH-DEPENDENT FLAVIN OXIDOREDUCTASE.  
 GN OrderedLocusNames=MYPU\_7720;  
 OS Mycoplasma pulmonis.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2107;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UAB CTIP;  
 RX MEDLINE=21267165; PubMed=11353084;  
 RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,  
 RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,  
 RA Blanchard A.;  
 RT "The complete genome sequence of the murine respiratory pathogen  
 RT Mycoplasma pulmonis";  
 RL Nucleic Acids Res. 29:2145-2153(2001).  
 DR EMBL; AL445565; CAC13945.1; -;  
 DR PIR; D90608; D90608.  
 DR MyPUList; MYPU\_7720; -;  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR003009; FMN enzyme.  
 DR InterPro; IPR001155; Oxidored FMN.  
 DR Pfam; PF00724; Oxidored FMN; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 398 AA; 45239 MW; 72A00D3FF61DBCC7 CRC64;  
 Query Match 96.4%; Score 27; DB 2; Length 398;  
 Best Local Similarity 83.3%; Pred. No. 5.1e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVEE 6  
 DB 243 GYSVEE 248  
 RESULT 54  
 NEUA STRA3 STANDARD; PRT; 413 AA.  
 AC Q53598; Q9S0S5;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE N-acetylneuraminic acid synthetase (EC 2.7.7.43) (CMP-N-  
 DE acetylneuraminic acid synthetase).  
 GN Name=neuA; OrderedLocusNames=gsb1233;  
 OS Streptococcus agalactiae (serotype III);  
 OC Streptococcus agalactiae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=216495, 1311;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COH31 / Serotype III;  
 RX MEDLINE=96228704; PubMed=8830246;  
 RA Haft R.F., Wessels M.R., Mebane M.F., Conaty N., Rubens C.E.;  
 RT "Characterization of cpsF and its product CMP-N-acetylneuraminic acid  
 RT synthetase, a group B streptococcal enzyme that can function in K1  
 RT capsular polysaccharide biosynthesis in Escherichia coli";  
 RL Mol. Microbiol. 19:555-563(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COH1 / Serotype III;  
 RX MEDLINE=20372630; PubMed=10913080;  
 RA Chaffin D.O., Beres S.B., Yim H.H., Rubens C.E.;  
 RT "The serotype of type Ia and III group B streptococci is determined by  
 RT the polymerase gene within the polycistronic capsule operon";  
 RL J. Bacteriol. 182:4466-4477(2000).  
 RN [3]

RP SEQUENCE FROM N.A.  
 RC STRAIN=NEW316 / Serotype III;  
 RX MEDLINE=22242508; PubMed=12354221;  
 RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Prangeul L.,  
 RA Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,  
 RA Kunst F.;  
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing  
 RT invasive neonatal disease";  
 RL Mol. Microbiol. 45:1499-1513(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OIL / Serotype Ia;  
 RX MEDLINE=99395021; PubMed=10464185;  
 RA Yamamoto S., Miyake K., Koike Y., Watanabe M., Machida Y., Ohta M.,  
 RA Iijima S.;  
 RT "Molecular characterization of type-specific capsular polysaccharide  
 RT biosynthesis genes of Streptococcus agalactiae type Ia";  
 RL J. Bacteriol. 181:5176-5184(1999).  
 CC -!- CARBOLYTIC ACTIVITY: CTP + N-acetylneuraminate = diphosphate + CMP-N-  
 CC acetylneuraminate.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the CMP-NeuNAc synthetase family.  
 CC -!- CAUTION: Was originally (Ref.1) called cpsF.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; U19899; AAB50271.1; -;  
 DR EMBL; AF163833; AAD53077.1; -;  
 DR EMBL; AL766849; CAD46892.1; -;  
 DR EMBL; AB028896; BAA82290.1; -;  
 DR PIR; S70912; S70912.  
 DR HSPP; Q57385; IREI.  
 DR SagalList; gbs1233; -;  
 DR InterPro; IPR003329; Cytidylyl trans.  
 DR Pfam; PF02348; CTP transf 3; 1\_  
 KW Bacterial capsule; Complete proteome; Exopolysaccharide synthesis;  
 KW Nucleotidyltransferase; Sialic acid; Transference.  
 FT CONFLICT 392 413 TOEANDLQSQYQLFVDEVKTL -> DSRG (in Ref. 1  
 FT and 2).  
 SQ SEQUENCE 413 AA; 47670 MW; 9089CD673ABD1FC8 CRC64;  
 Query Match 96.4%; Score 27; DB 1; Length 413;  
 Best Local Similarity 83.3%; Pred. No. 5.3e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GYXVEE 6  
 DB 306 GYSVEE 311  
 RESULT 55  
 NEUA STRA5 STANDARD; PRT; 413 AA.  
 AC Q9AFG9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE N-acetylneuraminic acid synthetase (EC 2.7.7.43) (CMP-N-  
 DE acetylneuraminic acid synthetase).  
 GN Name=neuA; OrderedLocusNames=SAG1158;  
 OS Streptococcus agalactiae (serotype V).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=216466;  
 RN [1]  
 RP SEQUENCE FROM N.A.



RC STRAIN-CNCTC 1/82 / Serotype V;  
RA McKinnon K., Chaffin D.O., Rubens C.E.;  
RT "Streptococcus agalactiae type V polysaccharide synthesis operon  
complete sequence."  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2603 V/R / Serotype V;  
RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;  
RA Tetelin H., Massignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,  
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,  
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,  
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,  
RA Radune D., Fedorova N.B., Scanlan D., Khouri H.W., Mulligan S.,  
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,  
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Matone D.,  
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,  
RA Praeger C.M.;  
RT "Complete genome sequence and comparative genomic analysis of an  
emerging human pathogen, serotype V Streptococcus agalactiae."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).  
CC -!- CATALYTIC ACTIVITY: CTP + N-acetylneuraminate + CMP-N-  
acylneuraminate.  
CC -!- SURCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the CMP-NeuNAc synthetase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF349539; AAK29664.1; -;  
DR EMBL: AE014244; AAN0040.1; -;  
DR HSSP: Q57385; IE21.  
DR TIGR: SAG1158; -;  
DR InterPro: IPR003329; Cytidylyl trans.  
DR Pfam: PF02348; CTP transf\_3; 1.  
KW Complete proteome; Nucleotidyltransferase; Sialic acid; Transferase.  
SQ SEQUENCE 413 AA; 47644 MW; 0AB5C0554F44172D CRC64;  
  
Query Match 96.4%; Score 27; DB 1; Length 413;  
Best Local Similarity 83.3%; Pred. No. 5.3e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVEE 6  
DB 306 GYSVEE 311  
  
RESULT 56  
Q9ALW4 PRELIMINARY; PRT; 413 AA.  
AC Q9ALW4  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE CMP-N-acetylneuraminic acid synthetase NeuA.  
GN Name=neuA;  
OS Streptococcus agalactiae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1311;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NT6;  
RA McKinnon K., Chaffin D.O., Rubens C.E.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF337958; AAK11673.1; -;  
DR HSSP: Q57385; IE21.  
DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.

DR InterPro: IPR003329; Cytidylyl trans.  
DR Pfam: PF02348; CTP transf\_3; 1. 0F4DA99B9DED97D37 CRC64;  
SQ SEQUENCE 413 AA; 47626 MW; 0F4DA99B9DED97D37 CRC64;  
  
Query Match 96.4%; Score 27; DB 2; Length 413;  
Best Local Similarity 83.3%; Pred. No. 5.3e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVEE 6  
DB 306 GYSVEE 311  
  
RESULT 57  
Q93TI0 PRELIMINARY; PRT; 413 AA.  
AC Q93TI0  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE CMP-N-acetylneuraminic acid synthetase NeuA.  
GN Name=neuA;  
OS Streptococcus agalactiae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1311;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CNCTC 1/82;  
RA McKinnon K., Chaffin D.O., Rubens C.E.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF355776; AAK43618.1; -;  
DR HSSP: Q57385; IE21.  
DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.  
DR InterPro: IPR003329; Cytidylyl trans.  
DR Pfam: PF02348; CTP transf\_3; 1.  
SQ SEQUENCE 413 AA; 47542 MW; 63A5C21C427DEF59 CRC64;  
  
Query Match 96.4%; Score 27; DB 2; Length 413;  
Best Local Similarity 83.3%; Pred. No. 5.3e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVEE 6  
DB 306 GYSVEE 311  
  
RESULT 58  
Q7WLQ2 PRELIMINARY; PRT; 431 AA.  
AC Q7WLQ2  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Portal protein.  
GN OrderedLocNames=BB1693;  
OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Alcaligenaceae; Bordetella.  
OX NCBI\_TaxID=518;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RB50 / ATCC BAA-588;  
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R., K.L.,  
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Cherevach I.,  
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Doggett J.,  
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,  
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,

```
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640442; CAE32190.1; -.
DR InterPro; IPR006944; Phage_portal.
DR InterPro; IPR006427; Portal_HK97.
DR Pfam; PF04860; Phage_portal; 1.
DR TIGRFAMs; TIGR01537; portal_HK97; 1.
KW Complete proteome.
SQ SEQUENCE 431 AA; 47542 MW; AE043AD93C0969 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 431;
Best Local Similarity 83.3%; Pred.No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 275 GYSVEE 280

RESULT 59
Q8AB03 PRELIMINARY; PRT; 447 AA.
ID Q8AB03
AC Q8AB03
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Dihydrolipoamide dehydrogenase.
GN OrderedLocusNames=BT0309;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
CC -!- CAPSULE ACTIVITY: Protein N(6)-(dihydrolipoyl) lysine + NAD(+) =
CC protein N(6)-(lipoyl)lysine + NADH.
CC -!- COFACTOR: Binds 1 PAD per subunit (By similarity).
CC -!- COFACTOR: FAD (By similarity).
CC -!- MISCELLANEOUS: The active site is a redox-active disulfide bond
CC (By similarity).
CC -!- SIMILARITY: Belongs to the class-I pyridine nucleotide-disulfide
CC oxidoreductase family.
DR EMBL; AE016927; AA075416.1; -.
DR HSSP; P31023; 1DXL.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0004148; F:dihydrolipoyl dehydrogenase activity; IEA.
DR GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR000815; Hg_reductase.
DR InterPro; IPR006258; Lipamide_dh.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR004099; Pyr_redox_dim.
DR Pfam; PF00070; Pyr_redox; 1.
DR Pfam; PF02852; Pyr_redox_dim; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00945; HGRDTASE.
DR PRINTS; PR00411; PNDRDTASEI.
DR PRINTS; PR00420; RNMNOXGNASE.
DR ProDom; PD000139; FAD_pyr_redox; 1.
DR TIGRFAMs; TIGR01350; lipamide_DH; 1.
```

```
DR PROSITE; PS00076; PYRIDINE REDOX 1; 1.
KW Complete proteome; FAD; Flavoprotein; NAD; Oxidoreductase;
KW Redox-active center.
SQ SEQUENCE 447 AA; 48578 MW; 1EF30F01A7665EC5 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 447;
Best Local Similarity 83.3%; Pred.No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 422 GYTVEE 427

RESULT 60
CGPB_FUSSO STANDARD; PRT; 457 AA.
ID CGPB_FUSSO
AC Q00858;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cutinase gene palindrome-binding protein (PBP).
OS Fusarium solani (subsp. pisi) (Nectria haematococca).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.
OX NCBI_TaxID=70791;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T-8;
RX MEDLINE=95263512; PubMed=7744822;
RA Li D., Kolattukudy P.E.;
RT "Cloning and expression of cDNA encoding a protein that binds a
RT palindromic promoter element essential for induction of fungal
RT cutinase by plant cutin.";
RL J. Biol. Chem. 270:11753-11756(1995).
CC -!- FUNCTION: Binds a palindromic promoter element essential for
CC induction of fungal cutinase gene.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 GATA-type zinc finger.
CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23722; AAA85727.1; -.
DR PIR; A57506; A57506.
DR HSSP; P17679; LGNF.
DR TRANSFAC; T02833; -.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR000679; ZnF_GATA.
DR Pfam; PF00320; GATA; 1.
DR Pfam; PF00989; PAS; 1.
DR SMART; SM00091; PAS; 1.
DR SMART; SM00401; ZnF_GATA; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
DR PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
DR PROSITE; PS00114; GATA_ZN_FINGER_2; 1.
DR PROSITE; PS00112; PAS; 1.
KW Activator; DNA-binding; Nuclear protein; Transcription regulation;
KW Zinc-finger.
FT DOMAIN 103 173 PAS.
FT ZN FING 402 427 GATA-type.
SQ SEQUENCE 457 AA; 49878 MW; 6B19E1452F518285 CRC64;

Query Match 96.4%; Score 27; DB 1; Length 457;
Best Local Similarity 83.3%; Pred.No. 5.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GYXVEE 6
Db      139 GYSVEE 144

RESULT 61
Q9CH26      PRELIMINARY;      PRT;      466 AA.
ID Q9CH26
AC Q9CH26;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Teichoic acid ABC transporter ATP binding protein.
GN Name=tagH; OrderedLocusNames=LL0915;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=11360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE006326; AA05013.1; -.
DR PIR; C86739; C86739.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF01476; LysM; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00257; LysM; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 466 AA; 5209 MW; 3A35DCD621656C48 CRC64;

Query Match      96.4%; Score 27; DB 2; Length 466;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
Db      421 GYSVEE 426

RESULT 62
Q8SVJ8      PRELIMINARY;      PRT;      528 AA.
ID Q8SVJ8
AC Q8SVJ8;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Similarity to HYPOTHETICAL GTP-BINDING PROTEINS OF THE GTPI/OBG
DE FAMILY.
GN Name=ECU05_0800;
OS Encephalitozoon cuniculi.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;

MEDLINE=21576510; PubMed=11719806;
RX Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RX Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
RX Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RX Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi."
RL Nature 414:450-453(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL590445; CAD26599.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR010674; NOG1.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF06858; NOG1; 1.
DR PRINTS; PR00326; GTP1_OBG.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW Hypothetical protein.
SQ SEQUENCE 528 AA; 61338 MW; E84804BA4EA5E9DE CRC64;

Query Match      96.4%; Score 27; DB 2; Length 528;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
Db      261 GYSVEE 266

RESULT 63
HUTU_STAAM      STANDARD;      PRT;      553 AA.
ID HUTU_STAAM
AC Q931GI;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Urocanate hydratase (EC 4.2.1.49) (Urocanase) (Imidazolonepropionate
DE hydrolase).
GN Name=hutU; OrderedLocusNames=SAV2331;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RX Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RX Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RX Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RX Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RX Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RX Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RX Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001)
CC -!- CATALYTIC ACTIVITY: 3-(5-oxo-4,5-dihydro-3-H-imidazol-4-
CC yl)propanoate = urocanate + H(2)O.
CC -!- COFACTOR: Binds 1 NAD per subunit (By similarity).
CC -!- PATHWAY: Histidine degradation; second step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the urocanase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; AP003365; BAB58493.1; -.
DR HAMAP; MF_00577; -.
DR InterPro; IPR010990; TFIIS conserved.
DR InterPro; IPR000193; Urocanase.
DR Pfam; PF01175; Urocanase; 1.
DR ProDom; PD025423; Urocanase; 1.
DR TIGRFAMs; TIGR01228; hutU; 1.
DR PROSITE; PS01233; UROCANASE; 1.
KW Complete proteome; Histidine metabolism; Lyase; NAD.
SQ SEQUENCE 553 AA; 60663 MW; 18CBB60A89747CE CRC64;

Query Match          96.4%; Score 27; DB 1; Length 553;
Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 270 GYSVEE 275

RESULT 64
HUTU_STAAN
ID HUTU_STAAN STANDARD; PRT; 553 AA.
AC Q99RU2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Urocanate hydratase (EC 4.2.1.49) (Urocanase) (Imidazolonepropionate
DE hydrolase).
GN Name=hutU; OrderedLocusNames=SA2122, MW2252;
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=159879, 196820;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Masumaru K., Maruyama A., Murakami H., Haseyama A.,
RA Mizutani-Ts Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiranatsu K.;
RA "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RA "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -|- CATALYTIC ACTIVITY: 3-(5-oxo-4,5-dihydro-3-H-imidazol-4-
CC yl)propanoate = urocanate + H(2)O.
CC -|- COFACTOR: Binds 1 NAD per subunit (By similarity).
CC -|- PATHWAY: Histidine degradation; second step.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the urocanase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; AP003366; BAB43423.1; -.
DR EMBL; AP004829; BAB96117.1; -.
DR PIR; F90032; F90032.
DR SWISS-2DPAGE; Q99RU2; STAAH.
DR HAMAP; MF_00577; -.
DR InterPro; IPR010990; TFIIS conserved.
DR InterPro; IPR000193; Urocanase.
DR Pfam; PF01175; Urocanase; 1.
DR ProDom; PD025423; Urocanase; 1.
DR TIGRFAMs; TIGR01228; hutU; 1.
DR PROSITE; PS01233; UROCANASE; 1.
KW Complete proteome; Histidine metabolism; Lyase; NAD.
SQ SEQUENCE 553 AA; 60632 MW; 19102455373CA8FB CRC64;

Query Match          96.4%; Score 27; DB 1; Length 553;
Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 270 GYSVEE 275

RESULT 65
Q6G6Y9 PRELIMINARY; PRT; 553 AA.
AC Q6G6Y9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Urocanate hydratase (EC 4.2.1.49).
GN ORFNames=SA52224;
OS Staphylococcus aureus subsp. aureus MSSA476.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=MSSA476;
RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RA "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571857; CAG44035.1; -.
DR InterPro; IPR010990; TFIIS conserved.
DR InterPro; IPR000193; Urocanase.
DR Pfam; PF01175; Urocanase; 1.
DR ProDom; PD025423; Urocanase; 1.
DR TIGRFAMs; TIGR01228; hutU; 1.
DR PROSITE; PS01233; UROCANASE; 1.
KW Lyase.
SQ SEQUENCE 553 AA; 60632 MW; 19102455373CA8FB CRC64;

Query Match          96.4%; Score 27; DB 2; Length 553;
Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 270 GYSVEE 275

RESULT 66
Q6GEA4 PRELIMINARY; PRT; 553 AA.
ID Q6GEA4
```

```

AC Q6GEA4;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DE Urocanate hydratase (EC 4.2.1.49)
GN Name=hutU; ORFNames=SAR2417;
OS Staphylococcus aureus subsp. aureus MRSA252.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282458;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MRSA252;
RA Holden M.T.G., Feil E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Armond D., Quail M.A., Rabinowitch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571856; CAG41397.1; -.
DR InterPro; IPR010990; TFIIS_Conserved.
DR InterPro; IPR000193; Urocanase.
DR Pfam; PF01175; Urocanase; 1.
DR ProDom; PD025423; Urocanase; 1.
DR TIGRFAMs; TIGR01228; hutU; 1.
DR PROSITE; PS01233; UROCANASE; 1.
KW Lyase.
SQ SEQUENCE 553 AA; 60642 MW; 9698B69BD6524723 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 553;
Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 270 GYSVEE 275

RESULT 67
Q84Q92 PRELIMINARY; PRT; 575 AA.
AC Q84Q92;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DE Hypothetical protein OJ1041F02.13.
GN Name=OJ1041F02.13;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Currie J.,
RA Collura K.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135206; AAP06856.1; -.
DR Gramene; O84Q92; -.
KW Hypothetical protein.
SQ SEQUENCE 575 AA; 60848 MW; 658C2643B3936684 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 575;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 270 GYSVEE 275

```

```

Db 499 GYTVEE 504

RESULT 68
Q7S720 PRELIMINARY; PRT; 645 AA.
AC Q7S720;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Related to cyclin B3.
GN Name=NCU01242.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Gaalagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaife D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gherre S.,
RA Kanal M., Kamyszelis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Piamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Navig D.O., Alex L.A., Mannhaupt G., Eboole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- SIMILARITY: Belongs to the cyclin family.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX0100267; EAA32372.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR InterPro; IPR004367; Cyclin_Cterm.
DR InterPro; IPR006671; Cyclin_N.
DR Pfam; PF02984; Cyclin_C; 1.
DR Pfam; PF00134; Cyclin_N; 1.
DR PROSITE; PS00292; CYCLINS; 1.
KW Cell cycle; Cell division; Cyclin.
SQ SEQUENCE 645 AA; 72415 MW; 856636EB2644EE22 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 645;
Best Local Similarity 83.3%; Pred. No. 8.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 484 GYSVEE 489

RESULT 69
Q8YT29 PRELIMINARY; PRT; 646 AA.
AC Q8YT29;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 26, Last annotation update)
DE Two-component hybrid sensor and regulator.
GN OrderedLocusNames=all2897;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriuchi M., Ishikawa A., Kawahima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shampo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213 (2001).
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL; AP003591; BAB74596.1; -.
DR PIR; AB2168; AB2168.
DR HSSP; P10958; 1DEW.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR011006; CheY_like.
DR InterPro; IPR005467; His_Kinase.
DR InterPro; IPR003661; His_kin_N.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR0001610; PAC.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR000700; PAS-assoc_C.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HisKA; 1.
DR Pfam; PF00785; PAC; 2.
DR Pfam; PF00989; PAS; 1.
DR Pfam; PF00072; Response_reg; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HisKA; 1.
DR SMART; SM00086; PAC; 2.
DR SMART; SM00091; PAC; 2.
DR SMART; SM00448; REC; 1.
DR SMART; SM00448; REC; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS01009; HIS_KIN; 1.
DR PROSITE; PS0113; PAC; 2.
DR PROSITE; PS0112; PAC; 2.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
DR Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferase.
SQ SEQUENCE 646 AA; 73599 MW; D773D0DFE12PFC43 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 646;
Best Local Similarity 83.3%; Pred. No. 8.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 314 GYSVEE 319

RESULT 70
Q9C2K2 PRELIMINARY; PRT; 653 AA.
AC Q9C2K2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Related to cyclin B3.
GN Name=3H10.100;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

```

```

OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
SEQUENCE FROM N.A.
RP Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RP German Neurospora genome project;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the cyclin family.
DR EMBL; AL513442; CAC28649.2; -.
DR HSSP; P20248; IFIN.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR InterPro; IPR006670; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR InterPro; IPR011028; Cyclin_like.
DR InterPro; IPR006671; Cyclin_N.
DR Pfam; PF02984; Cyclin_C; 1.
DR Pfam; PF00134; Cyclin_N; 1.
DR SMART; SM00385; CYCLIN; 2.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cell cycle; Cell division; Cyclin.
KW Cell cycle.
SQ SEQUENCE 653 AA; 73321 MW; 7250652D2C8B4996 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 653;
Best Local Similarity 83.3%; Pred. No. 8.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 492 GYSVEE 497

RESULT 71
Q88EF1 PRELIMINARY; PRT; 751 AA.
AC Q88EF1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PP4514;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfe E.K., Scanlan D., Tran K.,
RA Moazzaz A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Bisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808 (2002).
DR EMBL; AE016790; AAN70088.1; -.
DR TIGR; PP4514; -.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0045735; P:nutrient reservoir activity; IEA.
DR InterPro; IPR000184; Bac_surfAg_D15.
DR InterPro; IPR002641; Patatin.
DR Pfam; PF01103; Bac_surface_Ag; 1.
DR Pfam; PF01734; Patatin; 1.
KW Complete proteome; Hypothetical protein.

```

```
SQ SEQUENCE 751 AA; 83173 MW; CD140194E4CA6368 CRC64;
Query Match 96.4%; Score 27; DB 2; Length 751;
Best Local Similarity 83.3%; Pred. No. 9.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 97 GYSVEE 102

RESULT 72
GCR3 YEAST
ID GCR3 YEAST STANDARD; PRT; 861 AA.
AC P34160;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE GCR3 protein (STO1 protein) (SUT1 protein).
GN Name=GCR3; Synonyms=STO1, SUT1; OrderedLocusNames=YMR125W;
GN ORFNames=YMR564.07, YMR553.01;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92380925; PubMed=1512188;
RA Uemura H., Jigami Y.;
RT "GCR3 encodes an acidic protein that is required for expression of
glycolytic genes in Saccharomyces cerevisiae.";
RL J. Bacteriol. 174:5526-5532(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Pandit S., Sternglanz R.;
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Guo Z., Russo P., Sherman F.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lyne G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RN [5]
RL Nature 387:90-93(1997).
CC -!- FUNCTION: Required for expression of glycolytic genes. Has certain
characteristics of a transcriptional activator.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: SOME, TO HUMAN CBP80.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to a
frameshift in position 708.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10224; BAA01076.1; ALT SEQ.
DR EMBL; L07650; -; NOT ANNOTATED CDS.
DR EMBL; L27744; -; NOT ANNOTATED_CDS.
DR EMBL; Z49273; CAA89274.1; -.
DR EMBL; Z48622; CAA88550.1; -.
DR PIR; A44919; A44919.
DR GernOnline; 142794; -.
```

```
DR SGD; S0004732; STO1.
DR GO; GO:0000243; C:commitment complex; IPI.
DR GO; GO:0005846; C:snRNA cap binding complex; IDA.
DR GO; GO:0003729; F:mRNA binding; IPI.
DR GO; GO:0000398; P:nuclear mRNA splicing, via spliceosome; IPI.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR003890; IF_eIF4G.
DR Pfam; PF02854; MIF4G; 1_eIF4G.
DR DNA-binding; Nuclear protein.
KW DOMAIN 22 30
FT DOMAIN 774 801 Asp/Glu-rich (acidic).
FT DOMAIN 802 825 Arg/Lys-rich (basic).
FT CONFLICT 164 164 D -> V (in Ref. 3).
FT CONFLICT 633 633 R -> I (in Ref. 3).
FT CONFLICT 704 704 A -> R (in Ref. 3).
SQ SEQUENCE 861 AA; 100017 MW; EDD04907BDC9207D CRC64;
Query Match 96.4%; Score 27; DB 1; Length 861;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 568 GYTVEE 573

RESULT 73
Q8KQY6 PRELIMINARY; PRT; 877 AA.
AC Q8KQY6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein s038.
GN Name=s038;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22103115; PubMed=12107144;
RA Beaber J.W., Hochhut B., Waldor M.K.;
RT "Genomic and functional analyses of SXT, an integrating antibiotic
resistance gene transfer element derived from Vibrio cholerae.";
RL J. Bacteriol. 184:4259-4269(2002).
DR EMBL; AY055428; AAL59724.1; -.
KW Hypothetical protein.
SQ SEQUENCE 877 AA; 9489 MW; 8876B8C08B072CB4 CRC64;
Query Match 96.4%; Score 27; DB 2; Length 877;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 220 GYTVEE 225

RESULT 74
Q8ZS90 PRELIMINARY; PRT; 879 AA.
AC Q8ZS90;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cation-transporting ATPase.
GN OrderedLocusNames=alr7622;
OS Anabaena sp. (strain PCC 7120).
OG Plasmid pCC7120beta.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
```

RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iriuchichi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimp S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
RT cyanobacterium Anabaena sp. strain PCC 7120.";  
RL DNA Res. 8:205-213(2001).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type  
CC ATPases).  
DR EMBL; AP003602; BAB77265.1; -;  
DR PIR; AC2542; AC2542.  
DR HSP; O32221; IKOV.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:0016818; F:hydrolase activity, acting on acid anhydrid. . .; IEA.  
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.  
DR GO; GO:0046872; F:metal ion binding; IEA.  
DR GO; GO:0046873; F:metal ion transporter activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0030001; P:metal ion transport; IEA.  
DR GO; GO:0015992; P:proton transport; IEA.  
DR InterPro; IPR006416; ATPase-IB\_hvy.  
DR InterPro; IPR001757; ATPase\_E1-E2.  
DR InterPro; IPR005834; Dehal\_like\_hydro.  
DR InterPro; IPR008250; E1-E2\_ATPase\_reg.  
DR InterPro; IPR006121; HeavyMe\_transpt.  
DR InterPro; IPR006404; Heavy\_met\_ATPase.  
DR InterPro; IPR000695; H\_ATPase.  
DR InterPro; IPR006191; Metal\_bind.  
DR InterPro; IPR001969; Pept\_Asp\_AS.  
DR Pfam; PF0122; E1-E2\_ATPase; 1.  
DR Pfam; PF0403; HMA; 2.  
DR Pfam; PF0702; Hydrolase; 1.  
DR PRINTS; PR00119; CATATPASE.  
DR PRINTS; PR00120; HATPASE.  
DR TIGRFAMs; TIGR01512; ATPase-IB2\_Cd; 1.  
DR TIGRFAMs; TIGR01525; ATPase-IB\_hvy; 1.  
DR TIGRFAMs; TIGR01494; ATPase\_P-type; 2.  
DR PROSITE; PS00141; ASP\_PROTEASE; UNKNOWN\_1.  
DR PROSITE; PS00154; ATPASE\_E1\_E2; UNKNOWN\_1.  
DR PROSITE; PS01047; HMA\_1; 2.  
DR PROSITE; PS0846; HMA\_2; 2.  
DR ATP-binding; Complete proteome; Hydrolase; Phosphorylation; Plasmid;  
KW Transmembrane.  
SQ SEQUENCE 879 AA; 94953 MW; 887088B7CB8C3BD8 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 879;  
Best Local Similarity 83.3%; Pred.No. 1.1e+03;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 200 GTVEE 205

RESULT 75  
QB4A02  
ID QB4A02 PRELIMINARY; PRT; 885 AA.  
AC QB4A02;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein orf19.  
GN Name=orf19;  
OS Escherichia coli.  
OG Plasmid p1658/97.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zienkiewicz M., Kern-Zdanowicz I., Golebiewski M., Ceglowski P.;  
RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF550679; AA049574.1; -;  
DR KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 885 AA; 100704 MW; 826F3123E11E5D56 CRC64;  
Query Match 96.4%; Score 27; DB 2; Length 885;  
Best Local Similarity 83.3%; Pred.No. 1.1e+03;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GYXVEE 6  
|||  
Db 228 GTVEE 233  
Search completed: November 1, 2004, 21:34:27  
Job time : 173.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:23:25 ; Search time 28 Seconds  
(without alignments)  
14.211 Million cell updates/sec

Title: US-10-030-194A-5

Perfect score: 28

Sequence: 1 GYXVEE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	96.4	282	3	US-08-972-902-7
2	27	96.4	282	4	US-09-520-207-7
3	27	96.4	282	4	US-10-136-253-7
4	27	96.4	349	4	US-09-252-991A-27544
5	27	96.4	354	4	US-09-107-532A-5179
6	27	96.4	762	2	US-08-907-166-10
7	27	96.4	762	4	US-09-391-340-10
8	26	92.9	53	1	US-08-313-050-18
9	26	92.9	102	2	US-08-808-982-8
10	26	92.9	102	3	US-09-306-902A-8
11	26	92.9	107	4	US-09-173-300-13
12	26	92.9	115	4	US-09-173-300-13
13	26	92.9	118	4	US-09-585-173B-20
14	26	92.9	128	4	US-09-134-000C-3570
15	26	92.9	238	4	US-09-107-532A-6268
16	26	92.9	260	4	US-09-585-173B-22
17	26	92.9	307	4	US-09-173-300-15
18	26	92.9	326	3	US-09-564-805-231
19	26	92.9	345	1	US-08-183-214-12
20	26	92.9	357	3	US-09-134-001C-4405
21	26	92.9	358	4	US-09-710-279-18
22	26	92.9	362	3	US-09-134-001C-5403
23	26	92.9	369	4	US-09-543-681A-7314
24	26	92.9	378	4	US-09-270-767-59898
25	26	92.9	381	4	US-09-673-395A-363
26	26	92.9	412	4	US-09-543-681A-7895
27	26	92.9	425	4	US-09-485-529-6
28	26	92.9	460	3	US-09-134-001C-5322
29	26	92.9	532	3	US-09-117-853-2
30	26	92.9	532	4	US-09-911-154-2
31	26	92.9	532	4	US-09-485-529-2
32	26	92.9	532	4	US-09-911-514-2
33	26	92.9	623	4	US-09-485-529-7
34	26	92.9	630	4	US-09-485-529-8
35	26	92.9	698	4	US-09-489-847-200
36	26	92.9	705	4	US-09-270-767-4461
37	26	92.9	710	4	US-09-252-991A-24946
38	26	92.9	739	2	US-08-836-943-2
39	26	92.9	846	1	US-08-149-103-3
40	26	92.9	846	1	US-08-451-883-3
41	26	92.9	859	4	US-09-107-532A-6829
42	26	92.9	869	4	US-09-543-681A-4814
43	26	92.9	873	1	US-08-393-734-2
44	26	92.9	873	3	US-08-894-489-2
45	26	92.9	943	2	US-08-808-982-7
46	26	92.9	943	3	US-09-306-902A-7
47	26	92.9	945	4	US-10-140-002-146
48	26	92.9	1074	4	US-09-134-000C-6090
49	25	89.3	174	4	US-09-107-532A-5979
50	25	89.3	178	4	US-09-134-000C-3409
51	25	89.3	240	4	US-09-583-110-3479
52	25	89.3	311	3	US-09-564-805-230
53	25	89.3	311	4	US-09-393-858-20
54	25	89.3	311	4	US-10-190-279-20
55	25	89.3	346	4	US-09-252-991A-26418
56	25	89.3	522	4	US-09-489-039A-10041
57	25	89.3	869	4	US-09-489-039A-7727
58	25	89.3	889	4	US-09-252-991A-18702
59	25	89.3	912	3	US-08-943-768-2
60	25	89.3	912	4	US-09-865-960-2
61	24	85.7	13	2	US-08-407-804-16
62	24	85.7	13	3	US-09-124-807-16
63	24	85.7	60	4	US-09-513-999C-5989
64	24	85.7	64	4	US-09-621-976-6862
65	24	85.7	74	4	US-09-603-448-13
66	24	85.7	96	4	US-09-134-000C-6116
67	24	85.7	101	3	US-09-134-001C-3863
68	24	85.7	133	4	US-09-732-210-85
69	24	85.7	136	4	US-09-732-210-1505
70	24	85.7	143	4	US-09-710-279-1714
71	24	85.7	152	4	US-09-732-210-1504
72	24	85.7	194	4	US-09-252-991A-22894
73	24	85.7	200	4	US-09-134-000C-4202
74	24	85.7	264	4	US-09-710-279-776
75	24	85.7	268	3	US-09-134-001C-4729
76	24	85.7	284	4	US-09-252-991A-32820
77	24	85.7	315	4	US-09-386-653A-9
78	24	85.7	334	4	US-09-489-039A-9328
79	24	85.7	349	3	US-08-762-500-77
80	24	85.7	350	1	US-07-868-353A-14
81	24	85.7	350	2	US-08-407-804-23
82	24	85.7	350	3	US-09-124-807-23
83	24	85.7	351	4	US-09-107-532A-4693
84	24	85.7	371	4	US-09-134-000C-6089
85	24	85.7	523	4	US-09-134-000C-6006
86	24	85.7	528	4	US-09-107-532A-6340
87	24	85.7	532	4	US-09-543-681A-7614
88	24	85.7	627	4	US-09-403-667A-2
89	24	85.7	630	4	US-09-248-796A-19619
90	24	85.7	630	4	US-09-248-796A-19137
91	24	85.7	781	4	US-09-751-687-16
92	24	85.7	865	4	US-09-489-039A-9267
93	24	85.7	1095	4	US-09-710-279-2906
94	24	85.7	1572	4	US-09-543-681A-7609
95	24	85.7	1716	4	US-09-621-976-7003
96	23	82.1	54	4	US-09-270-767-33535
97	23	82.1	60	4	US-09-270-767-48752
98	23	82.1	60	4	US-09-107-532A-5286
99	23	82.1	70	4	US-08-912-272-9
100	23	82.1	77	3	US-08-912-272-9

Sequence 5322, Ap  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 7, Appli  
Sequence 8, Appli  
Sequence 200, App  
Sequence 44461, A  
Sequence 24946, A  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 6829, Ap  
Sequence 4814, Ap  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 7, Appli  
Sequence 146, App  
Sequence 6090, Ap  
Sequence 5979, Ap  
Sequence 3409, Ap  
Sequence 3479, Ap  
Sequence 230, App  
Sequence 20, Appli  
Sequence 20, Appli  
Sequence 26418, A  
Sequence 10041, A  
Sequence 7727, Ap  
Sequence 18702, A  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 16, Appli  
Sequence 16, Appli  
Sequence 5989, Ap  
Sequence 6862, Ap  
Sequence 13, Appli  
Sequence 6116, Ap  
Sequence 3863, Ap  
Sequence 85, Appli  
Sequence 1505, Ap  
Sequence 1714, Ap  
Sequence 1504, Ap  
Sequence 22894, A  
Sequence 4202, Ap  
Sequence 776, App  
Sequence 4729, Ap  
Sequence 32820, A  
Sequence 9, Appli  
Sequence 9328, Ap  
Sequence 77, Appli  
Sequence 14, Appli  
Sequence 23, Appli  
Sequence 23, Appli  
Sequence 4693, Ap  
Sequence 6089, Ap  
Sequence 6006, Ap  
Sequence 6340, Ap  
Sequence 7614, Ap  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 19619, A  
Sequence 19137, A  
Sequence 16, Appli  
Sequence 9267, Ap  
Sequence 2906, Ap  
Sequence 7609, Ap  
Sequence 7003, Ap  
Sequence 33535, A  
Sequence 48752, A  
Sequence 5286, Ap  
Sequence 9, Appli

101	23	82.1	77	3	US-09-026-039-9	Sequence 9, Appli	174	23	82.1	413	4	US-09-936-499-1	Sequence 1, Appli
102	23	82.1	84	4	US-09-248-796A-27018	Sequence 23017, A	175	23	82.1	413	4	US-09-936-499-2	Sequence 2, Appli
103	23	82.1	85	4	US-09-248-796A-23917	Sequence 27018, A	176	23	82.1	413	5	PCT-US91-00909-4	Sequence 4, Appli
104	23	82.1	103	4	US-09-732-210-763	Sequence 763, App	177	23	82.1	419	4	US-09-270-767-45469	Sequence 45469, A
105	23	82.1	105	4	US-09-732-210-777	Sequence 777, App	178	23	82.1	423	4	US-09-341-446B-9	Sequence 9, Appli
106	23	82.1	111	4	US-09-270-767-58884	Sequence 58884, A	179	23	82.1	440	4	US-09-198-452A-44	Sequence 44, Appli
107	23	82.1	125	4	US-09-252-991A-28602	Sequence 28602, A	180	23	82.1	448	2	US-08-884-072-1	Sequence 1, Appli
108	23	82.1	139	4	US-09-710-279-3044	Sequence 3044, Ap	181	23	82.1	448	2	US-09-212-168-1	Sequence 1, Appli
109	23	82.1	139	4	US-09-710-279-3302	Sequence 3302, Ap	182	23	82.1	448	4	US-09-409-096-4	Sequence 4, Appli
110	23	82.1	172	4	US-09-248-796A-14431	Sequence 14431, A	183	23	82.1	448	4	US-09-583-110-2755	Sequence 2755, Ap
111	23	82.1	181	4	US-09-543-681A-5968	Sequence 5968, Ap	184	23	82.1	469	4	US-09-107-532A-5936	Sequence 5936, Ap
112	23	82.1	183	4	US-09-252-991A-19168	Sequence 19168, A	185	23	82.1	493	4	US-09-411-628-10	Sequence 10, Appli
113	23	82.1	187	4	US-09-248-796A-19632	Sequence 19632, A	186	23	82.1	493	4	US-10-174-794-10	Sequence 10, Appli
114	23	82.1	195	4	US-09-134-000C-6158	Sequence 6158, Ap	187	23	82.1	514	3	US-09-066-047-3	Sequence 3, Appli
115	23	82.1	196	4	US-09-270-767-59006	Sequence 59006, A	188	23	82.1	523	4	US-09-252-991A-31596	Sequence 31596, A
116	23	82.1	211	3	US-09-162-184-34	Sequence 34, Appli	189	23	82.1	530	3	US-08-975-762-73	Sequence 73, Appli
117	23	82.1	211	3	US-09-489-777A-34	Sequence 34, Appli	190	23	82.1	530	3	US-09-295-028-73	Sequence 73, Appli
118	23	82.1	223	4	US-09-270-767-46327	Sequence 46327, A	191	23	82.1	530	3	US-09-180-852-2	Sequence 2, Appli
119	23	82.1	225	4	US-09-107-532A-6303	Sequence 6303, Ap	192	23	82.1	530	3	US-09-106-582-73	Sequence 73, Appli
120	23	82.1	228	3	US-09-134-001C-4373	Sequence 4373, Ap	193	23	82.1	530	3	US-09-199-637A-130	Sequence 130, App
121	23	82.1	230	4	US-09-252-991A-20963	Sequence 20963, A	194	23	82.1	530	4	US-09-159-469-73	Sequence 73, Appli
122	23	82.1	230	4	US-09-248-796A-16642	Sequence 16642, A	195	23	82.1	530	4	US-09-693-542-73	Sequence 73, Appli
123	23	82.1	237	4	US-09-489-039A-13697	Sequence 13697, A	196	23	82.1	537	4	US-09-252-991A-29592	Sequence 29592, A
124	23	82.1	254	4	US-09-583-110-5019	Sequence 5019, Ap	197	23	82.1	546	3	US-09-066-047-2	Sequence 2, Appli
125	23	82.1	257	4	US-09-679-279-7	Sequence 7, Appli	198	23	82.1	548	4	US-09-227-421-5	Sequence 5, Appli
126	23	82.1	264	4	US-09-886-319A-67	Sequence 67, Appli	199	23	82.1	548	4	US-09-479-855-5	Sequence 5, Appli
127	23	82.1	277	1	US-08-400-413-1	Sequence 1, Appli	200	23	82.1	555	4	US-09-272-421-2	Sequence 2, Appli
128	23	82.1	277	4	US-09-248-796A-20073	Sequence 20073, A	201	23	82.1	555	4	US-09-479-855-2	Sequence 2, Appli
129	23	82.1	302	3	US-09-282-305-14	Sequence 14, Appli	202	23	82.1	563	3	US-09-134-001C-2916	Sequence 2916, Ap
130	23	82.1	302	4	US-09-883-720-14	Sequence 14, Appli	203	23	82.1	590	3	US-08-975-762-64	Sequence 64, Appli
131	23	82.1	305	3	US-09-282-305-12	Sequence 12, Appli	204	23	82.1	590	3	US-09-295-028-64	Sequence 64, Appli
132	23	82.1	305	4	US-09-883-720-12	Sequence 12, Appli	205	23	82.1	590	3	US-09-106-582-64	Sequence 64, Appli
133	23	82.1	307	4	US-09-393-858-17	Sequence 17, Appli	206	23	82.1	590	4	US-09-159-469-64	Sequence 64, Appli
134	23	82.1	307	4	US-10-190-279-17	Sequence 17, Appli	207	23	82.1	590	4	US-09-693-542-64	Sequence 64, Appli
135	23	82.1	311	3	US-09-282-305-16	Sequence 16, Appli	208	23	82.1	628	4	US-09-107-532A-6437	Sequence 6437, Ap
136	23	82.1	311	4	US-09-883-720-16	Sequence 16, Appli	209	23	82.1	709	3	US-08-968-563-18	Sequence 18, Appli
137	23	82.1	317	4	US-09-543-681A-6261	Sequence 6261, Ap	210	23	82.1	709	3	US-08-969-683A-18	Sequence 18, Appli
138	23	82.1	324	4	US-09-248-796A-18761	Sequence 18761, A	211	23	82.1	709	3	US-09-297-928-15	Sequence 15, Appli
139	23	82.1	327	4	US-10-067-443-5	Sequence 5, Appli	212	23	82.1	717	4	US-09-134-000C-5833	Sequence 5833, Ap
140	23	82.1	328	3	US-09-180-827-7	Sequence 7, Appli	213	23	82.1	733	4	US-09-248-796A-16565	Sequence 16565, A
141	23	82.1	329	4	US-09-134-000C-6238	Sequence 6238, Ap	214	23	82.1	736	4	US-09-107-532A-6007	Sequence 6007, Ap
142	23	82.1	334	3	US-08-975-762-28	Sequence 28, Appli	215	23	82.1	750	6	5457037-3	Patent No. 5457037
143	23	82.1	334	3	US-08-821-324-28	Sequence 28, Appli	216	23	82.1	751	4	US-09-252-991A-33073	Patent No. 33073, A
144	23	82.1	334	3	US-09-295-028-28	Sequence 28, Appli	217	23	82.1	751	4	US-09-252-991A-33073	Patent No. 33073, A
145	23	82.1	334	3	US-09-106-582-28	Sequence 28, Appli	218	23	82.1	753	4	US-09-248-796A-19041	Sequence 19041, A
146	23	82.1	334	4	US-09-159-459-28	Sequence 28, Appli	219	23	82.1	772	2	US-08-410-784A-5	Sequence 5, Appli
147	23	82.1	334	4	US-09-693-542-28	Sequence 28, Appli	220	23	82.1	776	3	US-09-346-237-4	Sequence 4, Appli
148	23	82.1	335	4	US-09-312-283C-186	Sequence 186, App	221	23	82.1	776	3	US-09-346-237-7	Sequence 7, Appli
149	23	82.1	336	4	US-09-328-352-5838	Sequence 5838, Ap	222	23	82.1	846	1	US-08-149-103-4	Sequence 4, Appli
150	23	82.1	338	4	US-09-489-039A-10172	Sequence 10172, A	223	23	82.1	846	1	US-08-451-883-4	Sequence 4, Appli
151	23	82.1	338	4	US-09-270-767-43629	Sequence 43629, A	224	23	82.1	892	4	US-09-328-352-8164	Sequence 8164, Ap
152	23	82.1	341	4	US-09-134-000C-5524	Sequence 5524, Ap	225	23	82.1	1363	1	US-09-252-991A-21342	Sequence 21342, A
153	23	82.1	345	4	US-09-489-039A-8306	Sequence 8306, Ap	226	23	78.6	12	1	US-08-036-555B-164	Sequence 164, App
154	23	82.1	352	1	US-08-785-052-2	Sequence 2, Appli	227	22	78.6	12	1	US-08-469-569-164	Sequence 164, App
155	23	82.1	352	2	US-08-913-581-2	Sequence 2, Appli	228	22	78.6	12	1	US-08-249-322A-164	Sequence 164, App
156	23	82.1	352	4	US-09-495-406-22	Sequence 22, Appli	229	22	78.6	12	1	US-08-469-526A-164	Sequence 164, App
157	23	82.1	352	4	US-09-816-028A-36	Sequence 36, Appli	230	22	78.6	12	2	US-08-734-591A-164	Sequence 164, App
158	23	82.1	352	4	US-10-303-162-36	Sequence 36, Appli	231	22	78.6	12	2	US-08-469-660-164	Sequence 164, App
159	23	82.1	356	4	US-09-328-352-6258	Sequence 6258, Ap	232	22	78.6	12	3	US-08-470-335-164	Sequence 164, App
160	23	82.1	358	4	US-09-252-991A-20554	Sequence 20554, A	233	22	78.6	12	3	US-08-735-021-164	Sequence 164, App
161	23	82.1	361	3	US-09-134-001C-3862	Sequence 3862, Ap	234	22	78.6	12	3	US-08-734-664A-164	Sequence 164, App
162	23	82.1	362	1	US-08-118-270-15	Sequence 15, Appli	235	22	78.6	12	3	US-08-470-339-164	Sequence 164, App
163	23	82.1	362	5	PCT-US93-08528-15	Sequence 15, Appli	236	22	78.6	12	4	US-08-467-602-164	Sequence 164, App
164	23	82.1	378	4	US-09-325-932A-158	Sequence 158, App	237	22	78.6	12	5	PCT-US94-05083C-160	Sequence 160, App
165	23	82.1	392	4	US-09-489-039A-11256	Sequence 11256, A	238	22	78.6	12	5	PCT-US95-06848A-164	Sequence 164, App
166	23	82.1	399	2	US-08-742-621-3	Sequence 3, Appli	239	22	78.6	20	2	US-08-934-915-30	Sequence 30, Appli
167	23	82.1	399	2	US-08-750-134A-11	Sequence 11, Appli	240	22	78.6	20	2	US-08-934-915-135	Sequence 135, App
168	23	82.1	399	3	US-09-363-745-11	Sequence 11, Appli	241	22	78.6	35	3	US-09-146-950-7	Sequence 7, Appli
169	23	82.1	413	1	US-08-444-734A-5	Sequence 5, Appli	242	22	78.6	35	3	US-09-146-950-23	Sequence 23, Appli
170	23	82.1	413	1	US-08-087-772A-17	Sequence 17, Appli	243	22	78.6	51	3	US-07-946-180B-8	Sequence 8, Appli
171	23	82.1	413	1	US-08-467-568-12	Sequence 12, Appli	244	22	78.6	62	3	US-07-946-180B-9	Sequence 9, Appli
172	23	82.1	413	2	US-09-030-582-12	Sequence 12, Appli	245	22	78.6	68	4	US-09-270-767-45266	Sequence 45266, A
173	23	82.1	413	4	US-09-811-286-2	Sequence 2, Appli	246	22	78.6	68	2	US-08-454-557C-36	Sequence 36, Appli

247	22	78.6	68	2	US-08-340-426D-36	Sequence 36, Appl	320	22	78.6	209	5	PCT-US95-11869-2	Sequence 2, Appl
248	22	78.6	68	2	US-08-450-673C-36	Sequence 36, Appl	321	22	78.6	212	3	US-09-041-889-1	Sequence 1, Appl
249	22	78.6	68	5	PCT-US95-17111A-36	Sequence 36, Appl	322	22	78.6	212	3	US-08-837-058-1	Sequence 1, Appl
250	22	78.6	78	2	US-08-447-173A-58	Sequence 58, Appl	323	22	78.6	212	3	US-09-417-264-1	Sequence 1, Appl
251	22	78.6	78	2	US-08-447-173A-61	Sequence 61, Appl	324	22	78.6	213	1	US-09-609-324A-10	Sequence 10, Appl
252	22	78.6	81	4	US-09-583-110-3020	Sequence 3020, Ap	325	22	78.6	213	1	US-08-920-440B-10	Sequence 10, Appl
253	22	78.6	93	1	US-08-241-853-18	Sequence 18, Appl	326	22	78.6	213	3	US-09-173-492-10	Sequence 10, Appl
254	22	78.6	93	2	US-08-850-917-18	Sequence 18, Appl	327	22	78.6	213	3	US-09-173-133-10	Sequence 10, Appl
255	22	78.6	100	4	US-09-732-210-756	Sequence 756, App	328	22	78.6	213	4	US-09-580-236A-10	Sequence 10, Appl
256	22	78.6	103	3	US-09-041-889-39	Sequence 39, Appl	329	22	78.6	217	4	US-09-270-767-61796	Sequence 61796, A
257	22	78.6	103	4	US-09-417-264-39	Sequence 39, Appl	330	22	78.6	218	3	US-09-041-889-4	Sequence 4, Appl
258	22	78.6	104	1	US-09-609-324A-8	Sequence 8, Appl	331	22	78.6	218	3	US-08-837-058-4	Sequence 4, Appl
259	22	78.6	104	2	US-08-920-440B-8	Sequence 8, Appl	332	22	78.6	218	4	US-09-417-264-4	Sequence 4, Appl
260	22	78.6	104	3	US-09-173-492-8	Sequence 8, Appl	333	22	78.6	219	3	US-08-961-083-28	Sequence 28, Appl
261	22	78.6	104	3	US-09-173-133-8	Sequence 8, Appl	334	22	78.6	219	4	US-09-536-784-28	Sequence 28, Appl
262	22	78.6	104	3	US-09-165-533-8	Sequence 8, Appl	335	22	78.6	220	3	US-09-041-889-2	Sequence 2, Appl
263	22	78.6	104	4	US-09-580-236A-8	Sequence 8, Appl	336	22	78.6	220	3	US-08-837-058-2	Sequence 2, Appl
264	22	78.6	105	4	US-09-513-999C-5473	Sequence 5473, Ap	337	22	78.6	220	4	US-09-417-264-2	Sequence 2, Appl
265	22	78.6	108	4	US-09-538-092-368	Sequence 368, App	338	22	78.6	222	3	US-09-041-889-3	Sequence 3, Appl
266	22	78.6	112	3	US-09-142-469-5	Sequence 5, Appl	339	22	78.6	222	3	US-08-837-058-3	Sequence 3, Appl
267	22	78.6	113	4	US-09-134-000C-4812	Sequence 4812, Ap	340	22	78.6	222	4	US-09-417-264-3	Sequence 3, Appl
268	22	78.6	116	3	US-09-041-889-38	Sequence 38, Appl	341	22	78.6	224	3	US-08-944-483-35	Sequence 35, Appl
269	22	78.6	116	4	US-09-417-264-38	Sequence 38, Appl	342	22	78.6	225	4	US-09-328-352-6504	Sequence 6504, Ap
270	22	78.6	116	4	US-09-489-039A-11618	Sequence 11618, A	343	22	78.6	226	3	US-09-041-889-32	Sequence 32, Appl
271	22	78.6	122	4	US-09-270-767-32023	Sequence 32023, A	344	22	78.6	226	4	US-09-417-264-32	Sequence 32, Appl
272	22	78.6	131	4	US-09-134-000C-4619	Sequence 4619, Ap	345	22	78.6	227	4	US-09-248-796A-16910	Sequence 16910, A
273	22	78.6	140	4	US-09-134-000C-4154	Sequence 4154, Ap	346	22	78.6	229	4	US-09-252-991A-19104	Sequence 19104, A
274	22	78.6	141	4	US-09-107-532A-4712	Sequence 4712, Ap	347	22	78.6	231	4	US-09-583-110-4922	Sequence 4922, Ap
275	22	78.6	143	4	US-09-252-991A-32985	Sequence 32985, A	348	22	78.6	232	4	US-09-107-532A-6091	Sequence 6091, Ap
276	22	78.6	150	4	US-09-732-210-1368	Sequence 1368, Ap	349	22	78.6	238	4	US-09-328-352-6651	Sequence 6651, Ap
277	22	78.6	151	4	US-09-248-796A-24622	Sequence 24622, A	350	22	78.6	238	4	US-09-583-110-5286	Sequence 5286, Ap
278	22	78.6	154	4	US-09-538-092-1383	Sequence 1383, Ap	351	22	78.6	243	4	US-09-107-532A-4665	Sequence 4665, Ap
279	22	78.6	154	4	US-09-513-999C-8054	Sequence 8054, Ap	352	22	78.6	244	4	US-09-328-352-5728	Sequence 5728, Ap
280	22	78.6	155	3	US-09-146-950-4	Sequence 4, Appl	353	22	78.6	250	4	US-09-328-352-5728	Sequence 5728, Ap
281	22	78.6	158	3	US-09-041-889-40	Sequence 40, Appl	354	22	78.6	253	4	US-09-543-681A-5796	Sequence 5796, Ap
282	22	78.6	158	4	US-09-417-264-40	Sequence 40, Appl	355	22	78.6	254	4	US-09-252-991A-17980	Sequence 17980, A
283	22	78.6	159	3	US-09-146-950-20	Sequence 20, Appl	356	22	78.6	259	4	US-09-270-767-42085	Sequence 42085, A
284	22	78.6	161	4	US-09-252-991A-25087	Sequence 25087, A	357	22	78.6	260	4	US-09-107-532A-5350	Sequence 5350, Ap
285	22	78.6	161	4	US-09-523-323-56	Sequence 56, Appl	358	22	78.6	266	4	US-09-252-991A-21116	Sequence 21116, A
286	22	78.6	163	4	US-09-538-092-600	Sequence 600, App	359	22	78.6	277	4	US-09-107-532A-7213	Sequence 7213, Ap
287	22	78.6	164	4	US-09-134-000C-4743	Sequence 4743, Ap	360	22	78.6	280	4	US-09-270-767-40774	Sequence 40774, A
288	22	78.6	165	4	US-09-543-681A-6700	Sequence 6700, Ap	361	22	78.6	280	4	US-09-270-767-55990	Sequence 55990, A
289	22	78.6	182	4	US-09-615-192A-268	Sequence 268, App	362	22	78.6	282	4	US-09-252-991A-46679	Sequence 46679, A
290	22	78.6	184	1	US-09-609-324A-2	Sequence 2, Appl	363	22	78.6	283	3	US-08-807-151-1	Sequence 1, Appl
291	22	78.6	184	2	US-09-920-440B-2	Sequence 2, Appl	364	22	78.6	283	3	US-08-509-024-2	Sequence 2, Appl
292	22	78.6	184	3	US-09-173-492-2	Sequence 2, Appl	365	22	78.6	283	3	US-09-333-279-2	Sequence 2, Appl
293	22	78.6	184	3	US-09-173-133-2	Sequence 2, Appl	366	22	78.6	283	3	US-09-072-993C-2	Sequence 2, Appl
294	22	78.6	184	3	US-09-165-533-2	Sequence 2, Appl	367	22	78.6	283	3	US-09-478-937-1	Sequence 1, Appl
295	22	78.6	184	4	US-09-795-926-46	Sequence 46, Appl	368	22	78.6	283	4	US-09-631-780-2	Sequence 2, Appl
296	22	78.6	184	4	PCT-US95-12779-2	Sequence 2, Appl	369	22	78.6	283	5	PCT-US96-12374-2	Sequence 2, Appl
297	22	78.6	184	5	PCT-US95-15781-2	Sequence 2, Appl	370	22	78.6	283	5	PCT-US96-12374-2	Sequence 2, Appl
298	22	78.6	188	4	US-09-270-767-34220	Sequence 34220, A	371	22	78.6	295	4	US-09-071-252-11	Sequence 11, Appl
299	22	78.6	188	4	US-09-270-767-49437	Sequence 49437, A	372	22	78.6	296	3	US-09-134-001C-3652	Sequence 3652, Ap
300	22	78.6	188	4	US-09-270-767-49437	Sequence 49437, A	373	22	78.6	296	4	US-09-655-908-10	Sequence 10, Appl
301	22	78.6	192	4	US-09-107-532A-6267	Sequence 6267, Ap	374	22	78.6	296	4	US-09-710-279-2368	Sequence 2368, Ap
302	22	78.6	193	3	US-09-146-950-2	Sequence 2, Appl	375	22	78.6	297	2	US-08-859-106A-2	Sequence 2, Appl
303	22	78.6	195	4	US-09-248-796A-27805	Sequence 27805, A	376	22	78.6	297	3	US-09-142-469-2	Sequence 2, Appl
304	22	78.6	196	4	US-09-538-092-455	Sequence 455, App	377	22	78.6	297	4	US-09-402-664A-9	Sequence 9, Appl
305	22	78.6	197	3	US-09-146-950-18	Sequence 18, Appl	378	22	78.6	298	4	US-09-248-796A-18159	Sequence 18159, A
306	22	78.6	198	4	US-09-173-300-11	Sequence 11, Appl	379	22	78.6	300	3	US-09-134-001C-3422	Sequence 3422, Ap
307	22	78.6	202	4	US-09-252-991A-31316	Sequence 31316, A	380	22	78.6	305	4	US-09-866-061-9	Sequence 9, Appl
308	22	78.6	203	4	US-09-252-991A-18194	Sequence 18194, A	381	22	78.6	305	4	US-09-452-937A-44	Sequence 44, Appl
309	22	78.6	203	4	US-09-134-000C-5644	Sequence 5644, Ap	382	22	78.6	307	4	US-09-710-279-2122	Sequence 2122, Ap
310	22	78.6	204	4	US-09-270-767-32818	Sequence 32818, A	383	22	78.6	308	4	US-09-328-352-6112	Sequence 6112, Ap
311	22	78.6	204	4	US-09-270-767-48035	Sequence 48035, A	384	22	78.6	313	4	US-09-252-991A-32836	Sequence 32836, A
312	22	78.6	205	4	US-09-252-991A-20721	Sequence 20721, A	385	22	78.6	314	4	US-09-328-352-7094	Sequence 7094, Ap
313	22	78.6	206	3	US-09-041-889-6	Sequence 6, Appl	386	22	78.6	325	4	US-09-710-279-698	Sequence 698, App
314	22	78.6	206	3	US-08-837-058-6	Sequence 6, Appl	387	22	78.6	325	4	US-09-108-020-48	Sequence 48, Appl
315	22	78.6	206	4	US-09-417-264-6	Sequence 6, Appl	388	22	78.6	329	3	US-09-685-296-48	Sequence 48, Appl
316	22	78.6	209	1	US-08-455-001-2	Sequence 2, Appl	389	22	78.6	329	4	US-09-252-991A-23378	Sequence 23378, A
317	22	78.6	209	3	US-08-308-814-2	Sequence 2, Appl	390	22	78.6	330	4	US-09-252-991A-23378	Sequence 23378, A
318	22	78.6	209	4	US-09-214-631-6	Sequence 6, Appl	391	22	78.6	332	3	US-09-218-363-2	Sequence 2, Appl
319	22	78.6	209	4	US-08-393-462-2	Sequence 2, Appl	392	22	78.6	332	3	US-09-218-363-6	Sequence 6, Appl

393	22	78.6	332	3	US-09-218-363-9	Sequence 9, Appli	466	22	78.6	499	3	US-08-988-856B-4	Sequence 4, Appli
394	22	78.6	332	1	US-09-149-476-467	Sequence 467, App	467	22	78.6	499	4	US-09-843-845-2	Sequence 2, Appli
395	22	78.6	337	1	US-08-466-033-162	Sequence 162, App	468	22	78.6	499	4	US-09-843-845-4	Sequence 4, Appli
396	22	78.6	337	2	US-08-444-733-162	Sequence 162, App	469	22	78.6	499	4	US-09-917-254-60	Sequence 60, Appli
397	22	78.6	337	2	US-08-464-134-162	Sequence 162, App	470	22	78.6	499	5	PCT-US95-01806-2	Sequence 2, Appli
398	22	78.6	337	2	US-08-461-361-162	Sequence 162, App	471	22	78.6	499	5	PCT-US95-01806-4	Sequence 4, Appli
399	22	78.6	337	2	US-08-485-910-162	Sequence 162, App	472	22	78.6	506	4	US-09-489-039A-14248	Sequence 14248, A
400	22	78.6	339	3	US-08-688-988-2	Sequence 2, Appli	473	22	78.6	506	4	US-09-134-001C-4049	Sequence 4049, Ap
401	22	78.6	342	4	US-09-270-767-46228	Sequence 46228, A	474	22	78.6	509	4	US-09-252-991A-20470	Sequence 20470, A
402	22	78.6	345	3	US-09-134-001C-3774	Sequence 3774, A	475	22	78.6	516	1	US-08-638-911A-41	Sequence 41, Appli
403	22	78.6	354	1	US-07-868-353A-15	Sequence 15, Appli	476	22	78.6	523	4	US-09-252-991A-19841	Sequence 19841, A
404	22	78.6	354	2	US-08-407-804-24	Sequence 24, Appli	477	22	78.6	528	4	US-09-356-806-8	Sequence 8, Appli
405	22	78.6	354	3	US-09-124-807-24	Sequence 24, Appli	478	22	78.6	529	3	US-09-433-248A-6	Sequence 6, Appli
406	22	78.6	356	4	US-09-134-000C-6438	Sequence 6438, A	479	22	78.6	535	4	US-09-543-681A-4798	Sequence 4798, Ap
407	22	78.6	360	4	US-09-489-039A-13055	Sequence 13055, A	480	22	78.6	552	1	US-07-662-223-2	Sequence 2, Appli
408	22	78.6	362	1	US-08-183-214-2	Sequence 2, Appli	481	22	78.6	557	3	US-09-134-001C-4392	Sequence 4392, Ap
409	22	78.6	362	2	US-08-183-214-10	Sequence 10, Appli	482	22	78.6	560	4	US-09-248-796A-23013	Sequence 23013, A
410	22	78.6	365	2	US-08-472-666-4	Sequence 4, Appli	483	22	78.6	563	4	US-09-248-796A-20006	Sequence 20006, A
411	22	78.6	365	4	US-09-362-012A-2	Sequence 2, Appli	484	22	78.6	565	4	US-09-602-787A-616	Sequence 616, App
412	22	78.6	365	4	US-09-362-012A-5	Sequence 5, Appli	485	22	78.6	573	4	US-09-328-352-6016	Sequence 6016, Ap
413	22	78.6	365	5	PCT-US96-07615-4	Sequence 4, Appli	486	22	78.6	575	3	US-09-134-001C-3906	Sequence 3906, Ap
414	22	78.6	371	4	US-09-134-000C-5173	Sequence 5173, Ap	487	22	78.6	576	4	US-09-328-352-6971	Sequence 6971, Ap
415	22	78.6	375	4	US-09-270-767-31678	Sequence 31678, A	488	22	78.6	583	4	US-09-107-532A-6811	Sequence 6811, Ap
416	22	78.6	375	4	US-09-270-767-46895	Sequence 46895, A	489	22	78.6	584	1	US-08-461-822A-8	Sequence 8, Appli
417	22	78.6	377	3	US-09-023-023-2	Sequence 2, Appli	490	22	78.6	584	5	PCT-US95-0774A-8	Sequence 8, Appli
418	22	78.6	381	4	US-09-248-796A-14241	Sequence 14241, A	491	22	78.6	591	2	US-08-836-620A-17	Sequence 17, Appli
419	22	78.6	382	4	US-09-252-991A-29879	Sequence 29879, A	492	22	78.6	595	3	US-09-041-886-35	Sequence 35, Appli
420	22	78.6	382	4	US-09-543-681A-5932	Sequence 5932, Ap	493	22	78.6	595	3	US-08-764-870-12	Sequence 12, Appli
421	22	78.6	389	3	US-08-372-902-4	Sequence 4, Appli	494	22	78.6	595	3	US-08-980-115-12	Sequence 12, Appli
422	22	78.6	389	4	US-09-520-207-4	Sequence 4, Appli	495	22	78.6	595	4	US-08-453-998-2	Sequence 2, Appli
423	22	78.6	389	4	US-09-602-472A-57	Sequence 7, Appli	496	22	78.6	602	4	US-09-252-991A-24403	Sequence 24403, A
424	22	78.6	389	4	US-10-136-253-4	Sequence 4, Appli	497	22	78.6	617	4	US-09-489-039A-14321	Sequence 14321, A
425	22	78.6	390	4	US-09-252-991A-24517	Sequence 24517, A	498	22	78.6	637	4	US-09-489-039A-12521	Sequence 12521, A
426	22	78.6	398	4	US-09-489-039A-9541	Sequence 9541, A	499	22	78.6	637	4	US-09-118-276-6	Sequence 6, Appli
427	22	78.6	402	4	US-09-489-039A-7818	Sequence 7818, Ap	500	22	78.6	639	4	US-09-543-681A-6148	Sequence 6148, Ap
428	22	78.6	403	4	US-09-538-092-922	Sequence 922, App	501	22	78.6	639	4	US-09-198-452A-836	Sequence 836, App
429	22	78.6	419	3	US-08-509-024-7	Sequence 7, Appli	502	22	78.6	704	4	US-09-328-352-5537	Sequence 5537, App
430	22	78.6	419	3	US-09-333-279-7	Sequence 7, Appli	503	22	78.6	726	4	US-09-248-796A-17362	Sequence 17362, A
431	22	78.6	419	4	US-09-631-780-7	Sequence 7, Appli	504	22	78.6	732	4	US-09-134-000C-6359	Sequence 6359, Ap
432	22	78.6	421	4	US-09-252-991A-18560	Sequence 18560, A	505	22	78.6	737	4	US-09-602-787A-182	Sequence 182, App
433	22	78.6	421	4	US-09-248-796A-15893	Sequence 15893, A	506	22	78.6	740	4	US-09-538-092-1178	Sequence 1178, Ap
434	22	78.6	425	3	US-09-305-984-22	Sequence 22, Appli	507	22	78.6	761	4	US-09-585-858-28	Sequence 28, Appli
435	22	78.6	425	4	US-09-305-984-24	Sequence 24, Appli	508	22	78.6	810	4	US-09-489-039A-8577	Sequence 8577, Ap
436	22	78.6	425	4	US-09-073-541A-22	Sequence 22, Appli	509	22	78.6	860	4	US-09-489-039A-10975	Sequence 10975, A
437	22	78.6	425	4	US-09-493-940-22	Sequence 22, Appli	510	22	78.6	881	4	US-09-489-039A-12003	Sequence 12003, A
438	22	78.6	425	4	US-09-493-940-22	Sequence 24, Appli	511	22	78.6	885	4	US-09-538-092-1319	Sequence 1319, Ap
439	22	78.6	425	4	US-09-493-940-24	Sequence 24, Appli	512	22	78.6	885	4	US-09-248-796A-18062	Sequence 18062, A
440	22	78.6	425	4	US-09-563-110-5321	Sequence 5321, Ap	513	22	78.6	1060	4	US-09-538-092-191	Sequence 191, App
441	22	78.6	430	4	US-09-328-352-5931	Sequence 5931, Ap	514	22	78.6	1224	4	US-08-241-852-1	Sequence 2, Appli
442	22	78.6	431	4	US-09-489-039A-13702	Sequence 13702, A	515	22	78.6	1237	1	US-08-441-931-2	Sequence 2, Appli
443	22	78.6	440	4	US-09-248-796A-17060	Sequence 17060, A	516	22	78.6	1395	3	US-09-540-245A-15	Sequence 15, Appli
444	22	78.6	448	4	US-09-583-110-4489	Sequence 4489, Ap	517	22	78.6	1395	3	US-09-466-033-15	Sequence 15, Appli
445	22	78.6	452	1	US-08-317-880-2	Sequence 2, Appli	518	22	78.6	2873	1	US-08-638-911A-2	Sequence 2, Appli
446	22	78.6	452	2	US-08-782-396-2	Sequence 2, Appli	519	22	78.6	2873	1	US-08-444-733-15	Sequence 15, Appli
447	22	78.6	459	4	US-09-489-039A-7416	Sequence 7416, Ap	520	22	78.6	2873	2	US-08-464-134-15	Sequence 15, Appli
448	22	78.6	460	1	US-08-403-866-3	Sequence 3, Appli	521	22	78.6	2873	2	US-08-461-361-15	Sequence 15, Appli
449	22	78.6	464	3	US-09-426-072-2	Sequence 2, Appli	522	22	78.6	2873	2	US-08-485-910-15	Sequence 15, Appli
450	22	78.6	466	2	US-08-955-138-7	Sequence 7, Appli	523	22	78.6	2873	5	PCT-US95-06266-15	Sequence 15, Appli
451	22	78.6	476	2	US-08-828-488-3	Sequence 3, Appli	524	22	78.6	2905	4	US-08-469-260A-401	Sequence 401, App
452	22	78.6	476	3	US-09-299-689A-3	Sequence 3, Appli	525	22	78.6	2905	4	US-08-488-446-401	Sequence 401, App
453	22	78.6	476	4	US-09-489-039A-8120	Sequence 8120, Ap	526	22	78.6	2905	4	US-08-467-344A-401	Sequence 401, App
454	22	78.6	476	4	US-09-907-794A-164	Sequence 164, App	527	22	78.6	2905	4	US-08-424-550B-401	Sequence 401, App
455	22	78.6	476	4	US-09-905-125A-164	Sequence 164, App	528	22	78.6	2910	1	US-08-466-033-183	Sequence 183, App
456	22	78.6	476	4	US-09-902-775A-164	Sequence 164, App	529	22	78.6	2910	2	US-08-444-733-183	Sequence 183, App
457	22	78.6	476	4	US-09-906-700-164	Sequence 164, App	530	22	78.6	2910	2	US-08-464-134-183	Sequence 183, App
458	22	78.6	476	4	US-09-903-603A-164	Sequence 164, App	531	22	78.6	2910	2	US-08-461-361-183	Sequence 183, App
459	22	78.6	477	3	US-08-828-488-1	Sequence 1, Appli	532	22	78.6	2910	2	US-08-485-910-183	Sequence 183, App
460	22	78.6	477	3	US-09-299-689A-1	Sequence 1, Appli	533	22	78.6	2910	5	PCT-US95-06266-157	Sequence 157, App
461	22	78.6	489	4	US-09-107-532A-4711	Sequence 4711, Ap	534	22	78.6	3892	4	US-09-328-352-5503	Sequence 5503, Ap
462	22	78.6	492	3	US-09-342-749-2	Sequence 2, Appli	535	21	75.0	7	4	US-09-535-852-1127	Sequence 1127, Ap
463	22	78.6	492	4	US-09-691-840-2	Sequence 2, Appli	536	21	75.0	8	4	US-09-535-852-1132	Sequence 1132, Ap
464	22	78.6	496	4	US-08-622-191-1	Sequence 1, Appli	537	21	75.0	9	4	US-09-535-852-1137	Sequence 1137, Ap
465	22	78.6	499	3	US-08-988-856B-2	Sequence 2, Appli	538	21	75.0	13	2	US-08-407-804-19	Sequence 19, Appli

539	21	75.0	13	3	US-09-124-807-19	Sequence 19, Appl	612	21	75.0	91	4	US-09-220-527-89	Sequence 89, Appl
540	21	75.0	16	3	US-08-931-858E-94	Sequence 94, Appl	613	21	75.0	91	4	US-09-220-407-83	Sequence 83, Appl
541	21	75.0	16	3	US-08-981-739-94	Sequence 94, Appl	614	21	75.0	91	4	US-09-220-407-89	Sequence 89, Appl
542	21	75.0	16	4	US-09-128-026-94	Sequence 94, Appl	615	21	75.0	92	4	US-09-513-999C-5219	Sequence 5219, Ap
543	21	75.0	16	4	US-09-220-616-94	Sequence 94, Appl	616	21	75.0	94	2	US-08-465-380-22	Sequence 22, Appl
544	21	75.0	16	4	US-09-220-527-94	Sequence 94, Appl	617	21	75.0	94	2	US-08-480-478-51	Sequence 51, Appl
545	21	75.0	16	4	US-09-220-407-94	Sequence 94, Appl	618	21	75.0	94	2	US-08-486-397-22	Sequence 22, Appl
546	21	75.0	16	4	US-09-775-743C-15	Sequence 15, Appl	619	21	75.0	94	2	US-08-486-399-22	Sequence 22, Appl
547	21	75.0	18	3	US-08-825-852-66	Sequence 66, Appl	620	21	75.0	94	2	US-08-461-965-22	Sequence 22, Appl
548	21	75.0	18	3	US-09-052-888-67	Sequence 67, Appl	621	21	75.0	94	2	US-08-326-110A-51	Sequence 51, Appl
549	21	75.0	18	4	US-09-723-890-67	Sequence 67, Appl	622	21	75.0	94	2	US-08-634-641-22	Sequence 22, Appl
550	21	75.0	18	4	US-09-723-901-67	Sequence 67, Appl	623	21	75.0	94	2	US-09-249-471-22	Sequence 22, Appl
551	21	75.0	18	4	US-09-723-547-67	Sequence 67, Appl	624	21	75.0	94	2	US-09-249-472-22	Sequence 22, Appl
552	21	75.0	18	4	US-09-724-127-67	Sequence 67, Appl	625	21	75.0	94	2	US-08-249-451-22	Sequence 22, Appl
553	21	75.0	18	4	US-09-723-931-67	Sequence 67, Appl	626	21	75.0	94	2	US-08-809-455-22	Sequence 22, Appl
554	21	75.0	18	4	US-09-723-873-67	Sequence 67, Appl	627	21	75.0	94	2	US-09-249-461-22	Sequence 22, Appl
555	21	75.0	18	4	US-09-724-114-67	Sequence 67, Appl	628	21	75.0	94	2	US-09-249-448-22	Sequence 22, Appl
556	21	75.0	18	4	US-09-723-913-67	Sequence 67, Appl	629	21	75.0	94	2	US-09-249-473-22	Sequence 22, Appl
557	21	75.0	18	4	US-09-723-912-67	Sequence 67, Appl	630	21	75.0	94	2	US-09-270-767-56851	Sequence 56851, A
558	21	75.0	18	4	US-09-724-095-67	Sequence 67, Appl	631	21	75.0	94	2	US-09-270-767-59680	Sequence 59680, A
559	21	75.0	18	4	US-09-724-157-67	Sequence 67, Appl	632	21	75.0	94	2	US-09-270-767-59982	Sequence 59982, A
560	21	75.0	18	4	US-09-724-062-67	Sequence 67, Appl	633	21	75.0	94	2	US-09-270-767-61758	Sequence 61758, A
561	21	75.0	18	4	US-09-724-065-67	Sequence 67, Appl	634	21	75.0	95	3	US-09-374-135-8	Sequence 8, Appl
562	21	75.0	21	3	US-09-399-494-16	Sequence 16, Appl	635	21	75.0	95	3	US-08-931-858E-80	Sequence 80, Appl
563	21	75.0	21	3	US-09-399-494-17	Sequence 17, Appl	636	21	75.0	95	3	US-08-931-858E-141	Sequence 141, App
564	21	75.0	21	3	US-09-399-494-18	Sequence 18, Appl	637	21	75.0	95	3	US-08-931-858E-187	Sequence 187, App
565	21	75.0	32	4	US-09-230-041-11	Sequence 11, Appl	638	21	75.0	96	3	US-08-931-858E-198	Sequence 198, App
566	21	75.0	47	4	US-09-157-689-21	Sequence 21, Appl	639	21	75.0	96	3	US-08-931-858E-221	Sequence 221, App
567	21	75.0	47	4	US-09-157-689-22	Sequence 22, Appl	640	21	75.0	96	3	US-08-981-739-80	Sequence 80, Appl
568	21	75.0	47	4	US-08-447-398-21	Sequence 21, Appl	641	21	75.0	96	3	US-08-981-739-141	Sequence 141, App
569	21	75.0	47	4	US-08-447-398-22	Sequence 22, Appl	642	21	75.0	96	3	US-09-220-528-15	Sequence 15, Appl
570	21	75.0	63	4	US-09-270-767-57725	Sequence 57725, A	643	21	75.0	96	4	US-09-128-026-80	Sequence 80, Appl
571	21	75.0	64	4	US-09-248-796A-23545	Sequence 23545, A	644	21	75.0	96	4	US-09-128-026-141	Sequence 141, App
572	21	75.0	72	4	US-09-270-767-57855	Sequence 57855, A	645	21	75.0	96	4	US-09-220-616-80	Sequence 80, Appl
573	21	75.0	75	2	US-08-465-380-57	Sequence 57, Appl	646	21	75.0	96	4	US-09-220-616-141	Sequence 141, App
574	21	75.0	75	2	US-08-465-380-57	Sequence 57, Appl	647	21	75.0	96	4	US-09-220-527-80	Sequence 80, Appl
575	21	75.0	75	2	US-08-465-380-57	Sequence 57, Appl	648	21	75.0	96	4	US-09-220-527-141	Sequence 141, App
576	21	75.0	75	2	US-08-461-965-57	Sequence 57, Appl	649	21	75.0	96	4	US-09-220-407-80	Sequence 80, Appl
577	21	75.0	75	2	US-08-634-641-57	Sequence 57, Appl	650	21	75.0	96	4	US-09-220-407-141	Sequence 141, App
578	21	75.0	75	3	US-09-249-471-57	Sequence 57, Appl	651	21	75.0	96	4	US-09-220-407-187	Sequence 187, App
579	21	75.0	75	3	US-09-249-472-57	Sequence 57, Appl	652	21	75.0	96	4	US-09-220-407-198	Sequence 198, App
580	21	75.0	75	3	US-09-249-451-57	Sequence 57, Appl	653	21	75.0	96	4	US-09-220-407-221	Sequence 221, App
581	21	75.0	75	3	US-08-803-455-57	Sequence 57, Appl	654	21	75.0	96	4	US-09-220-407-221	Sequence 221, App
582	21	75.0	75	3	US-09-249-461-57	Sequence 57, Appl	655	21	75.0	97	4	US-09-489-039A-10490	Sequence 10490, A
583	21	75.0	75	3	US-09-249-448-57	Sequence 57, Appl	656	21	75.0	97	4	US-09-489-039A-13415	Sequence 13415, A
584	21	75.0	75	4	US-09-249-473-57	Sequence 57, Appl	657	21	75.0	104	3	US-08-858-207A-328	Sequence 328, App
585	21	75.0	80	4	US-09-270-767-45353	Sequence 45353, A	658	21	75.0	107	4	US-09-248-796A-18206	Sequence 18206, A
586	21	75.0	82	4	US-09-513-999C-4978	Sequence 4978, A	659	21	75.0	109	4	US-09-134-000C-5763	Sequence 5763, Ap
587	21	75.0	85	4	US-09-270-767-61109	Sequence 61109, A	660	21	75.0	111	4	US-09-513-999C-4203	Sequence 4203, Ap
588	21	75.0	89	3	US-08-931-858E-79	Sequence 79, Appl	661	21	75.0	111	4	Sequence 2957, Ap	
589	21	75.0	89	3	US-08-931-858E-82	Sequence 82, Appl	662	21	75.0	117	4	Sequence 29584, A	
590	21	75.0	89	3	US-08-931-858E-223	Sequence 223, App	663	21	75.0	117	4	Sequence 32171, A	
591	21	75.0	89	3	US-08-981-739-79	Sequence 79, Appl	664	21	75.0	117	4	Sequence 41066, A	
592	21	75.0	89	3	US-08-981-739-82	Sequence 82, Appl	665	21	75.0	117	4	Sequence 56282, A	
593	21	75.0	89	3	US-09-220-528-18	Sequence 18, Appl	666	21	75.0	117	4	Sequence 12, Appl	
594	21	75.0	89	4	US-09-128-026-79	Sequence 79, Appl	667	21	75.0	118	4	Sequence 4979, Ap	
595	21	75.0	89	4	US-09-128-026-82	Sequence 82, Appl	668	21	75.0	119	4	Sequence 58586, A	
596	21	75.0	89	4	US-09-220-616-79	Sequence 79, Appl	669	21	75.0	120	4	Sequence 32195, A	
597	21	75.0	89	4	US-09-220-616-82	Sequence 82, Appl	670	21	75.0	120	4	Sequence 47412, A	
598	21	75.0	89	4	US-09-220-527-79	Sequence 79, Appl	671	21	75.0	122	4	Sequence 11614, A	
599	21	75.0	89	4	US-09-220-527-82	Sequence 82, Appl	672	21	75.0	123	3	Sequence 3259, Ap	
600	21	75.0	89	4	US-09-220-407-82	Sequence 79, Appl	673	21	75.0	124	4	Sequence 5056, Ap	
601	21	75.0	89	4	US-09-220-407-82	Sequence 82, Appl	674	21	75.0	126	4	Sequence 4900, Ap	
602	21	75.0	89	4	US-09-220-407-223	Sequence 223, App	675	21	75.0	126	4	Sequence 5175, Ap	
603	21	75.0	91	3	US-08-931-858E-83	Sequence 83, Appl	676	21	75.0	127	4	Sequence 5037, Ap	
604	21	75.0	91	3	US-08-931-858E-89	Sequence 89, Appl	677	21	75.0	128	2	Sequence 3, Appl	
605	21	75.0	91	3	US-08-981-739-83	Sequence 83, Appl	678	21	75.0	128	3	Sequence 4265, Ap	
606	21	75.0	91	3	US-08-981-739-89	Sequence 89, Appl	679	21	75.0	131	4	Sequence 125, App	
607	21	75.0	91	4	US-09-128-026-83	Sequence 83, Appl	680	21	75.0	132	4	Sequence 132, App	
608	21	75.0	91	4	US-09-128-026-89	Sequence 89, Appl	681	21	75.0	133	3	Sequence 132, App	
609	21	75.0	91	4	US-09-220-616-83	Sequence 83, Appl	682	21	75.0	133	3	Sequence 132, App	
610	21	75.0	91	4	US-09-220-616-89	Sequence 89, Appl	683	21	75.0	134	3	Sequence 81, Appl	
611	21	75.0	91	4	US-09-220-527-83	Sequence 83, Appl	684	21	75.0	134	4	Sequence 81, Appl	

685	21	75.0	134	4	US-09-220-527-81	Sequence 81, Appl	758	21	75.0	206	4	US-09-107-532A-6492	Sequence 6492, Ap
686	21	75.0	134	4	US-09-270-767-32905	Sequence 32905, A	759	21	75.0	208	1	US-07-935-309-2	Sequence 2, Appl
687	21	75.0	134	4	US-09-270-767-41015	Sequence 41015, A	760	21	75.0	208	1	US-08-039-364-2	Sequence 2, Appl
688	21	75.0	134	4	US-09-270-767-56231	Sequence 56231, A	761	21	75.0	208	3	US-08-718-904-5	Sequence 5, Appl
689	21	75.0	135	4	US-09-328-352-4505	Sequence 4505, Ap	762	21	75.0	208	3	US-08-718-904-7	Sequence 7, Appl
690	21	75.0	141	4	US-09-513-999C-5760	Sequence 5760, Ap	763	21	75.0	208	3	US-09-094-103-4	Sequence 4, Appl
691	21	75.0	142	3	US-08-931-858E-111	Sequence 111, App	764	21	75.0	208	3	US-09-181-974-2	Sequence 2, Appl
692	21	75.0	142	3	US-08-981-739-111	Sequence 111, App	765	21	75.0	208	3	US-09-158-710-2	Sequence 2, Appl
693	21	75.0	142	4	US-09-128-026-111	Sequence 111, App	766	21	75.0	208	3	US-09-158-710-2	Sequence 2, Appl
694	21	75.0	142	4	US-09-328-352-5761	Sequence 5761, Ap	767	21	75.0	208	3	US-09-518-950-2	Sequence 2, Appl
695	21	75.0	142	4	US-09-220-616-111	Sequence 111, App	768	21	75.0	208	4	US-09-449-249-5	Sequence 5, Appl
696	21	75.0	142	4	US-09-220-527-111	Sequence 111, App	769	21	75.0	208	4	US-09-449-249-7	Sequence 7, Appl
697	21	75.0	142	4	US-09-270-767-45594	Sequence 45594, A	770	21	75.0	208	4	US-10-138-158-18	Sequence 18, Appl
698	21	75.0	142	4	US-09-270-767-45594	Sequence 45594, A	771	21	75.0	208	4	US-09-248-796A-18276	Sequence 18276, A
699	21	75.0	143	3	US-08-858-207A-441	Sequence 441, App	772	21	75.0	211	1	US-07-915-934-4	Sequence 4, Appl
700	21	75.0	143	4	US-09-828-523A-98	Sequence 98, Appl	773	21	75.0	211	1	US-08-325-743-4	Sequence 4, Appl
701	21	75.0	144	4	US-09-828-523A-99	Sequence 99, Appl	774	21	75.0	211	4	US-09-248-796A-18236	Sequence 4, Appl
702	21	75.0	145	4	US-09-107-532A-3705	Sequence 3705, Ap	775	21	75.0	212	4	US-08-486-049-4	Sequence 4, Appl
703	21	75.0	149	4	US-09-590-101A-2	Sequence 2, Appl	776	21	75.0	212	4	US-09-489-039A-13107	Sequence 13107, A
704	21	75.0	150	4	US-09-543-681A-6017	Sequence 6017, Ap	777	21	75.0	213	4	US-09-252-991A-23391	Sequence 23391, A
705	21	75.0	155	3	US-09-203-716-4	Sequence 4, Appl	778	21	75.0	215	3	US-08-914-375C-29	Sequence 29, Appl
706	21	75.0	155	3	US-09-684-254-4	Sequence 4, Appl	779	21	75.0	215	4	US-09-328-352-4289	Sequence 4289, Ap
707	21	75.0	155	4	US-09-409-926-5	Sequence 5, Appl	780	21	75.0	216	4	US-09-489-039A-9538	Sequence 9538, Ap
708	21	75.0	155	4	US-09-270-767-33449	Sequence 33449, A	781	21	75.0	217	4	US-09-328-352-5784	Sequence 5784, Ap
709	21	75.0	155	4	US-09-270-767-48666	Sequence 48666, A	782	21	75.0	223	1	US-07-708-885B-3	Sequence 3, Appl
710	21	75.0	156	3	US-08-931-858E-185	Sequence 185, App	783	21	75.0	223	1	US-07-714-386-3	Sequence 3, Appl
711	21	75.0	156	3	US-08-931-858E-196	Sequence 196, App	784	21	75.0	223	1	US-07-708-888A-3	Sequence 3, Appl
712	21	75.0	156	3	US-08-931-858E-217	Sequence 217, App	785	21	75.0	223	4	US-09-248-796A-27107	Sequence 27107, A
713	21	75.0	156	4	US-09-347-613C-36	Sequence 36, Appl	786	21	75.0	224	2	US-08-766-982-13	Sequence 13, Appl
714	21	75.0	156	4	US-09-220-407-185	Sequence 185, App	787	21	75.0	224	3	US-08-944-483-34	Sequence 34, Appl
715	21	75.0	156	4	US-09-220-407-196	Sequence 196, App	788	21	75.0	224	3	US-08-944-483-36	Sequence 36, Appl
716	21	75.0	156	4	US-09-220-407-217	Sequence 217, App	789	21	75.0	224	3	US-09-296-219-13	Sequence 13, Appl
717	21	75.0	156	4	US-09-662-183A-36	Sequence 36, Appl	790	21	75.0	224	4	US-08-963-901-6	Sequence 6, Appl
718	21	75.0	157	4	US-09-270-767-36537	Sequence 36537, A	791	21	75.0	224	4	US-09-710-279-496	Sequence 496, App
719	21	75.0	161	4	US-09-270-767-51754	Sequence 51754, A	792	21	75.0	224	4	US-09-710-279-1366	Sequence 1366, Ap
720	21	75.0	165	4	US-09-489-039A-6820	Sequence 6820, Ap	793	21	75.0	225	4	US-09-027-337-5	Sequence 5, Appl
721	21	75.0	165	4	US-09-489-039A-14030	Sequence 14030, Ap	794	21	75.0	225	4	US-09-644-600-5	Sequence 5, Appl
722	21	75.0	165	4	US-09-270-767-44255	Sequence 44255, A	795	21	75.0	225	4	US-09-654-600A-5	Sequence 5, Appl
723	21	75.0	166	4	US-09-489-039A-12565	Sequence 12565, A	796	21	75.0	227	4	US-09-270-767-31967	Sequence 31967, A
724	21	75.0	167	4	US-09-252-991A-23665	Sequence 23665, A	797	21	75.0	227	4	US-09-270-767-47184	Sequence 47184, A
725	21	75.0	168	4	US-09-252-991A-25254	Sequence 25254, A	798	21	75.0	228	3	US-09-134-001C-5384	Sequence 5384, Ap
726	21	75.0	171	4	US-09-543-681A-8055	Sequence 8055, Ap	799	21	75.0	229	3	US-09-078-691-4	Sequence 4, Appl
727	21	75.0	173	4	US-09-252-991A-21967	Sequence 21967, A	800	21	75.0	229	4	US-09-252-991A-31789	Sequence 31789, A
728	21	75.0	173	4	US-09-583-110-4353	Sequence 4353, Ap	801	21	75.0	229	4	US-09-583-110-3966	Sequence 3966, Ap
729	21	75.0	174	4	US-09-270-767-45254	Sequence 45254, A	802	21	75.0	229	4	US-09-583-110-5283	Sequence 5283, Ap
730	21	75.0	180	4	US-09-252-991A-26385	Sequence 26385, A	803	21	75.0	231	2	US-08-286-819A-12	Sequence 12, Appl
731	21	75.0	181	4	US-09-854-864-5	Sequence 5, Appl	804	21	75.0	231	2	US-08-286-819A-43	Sequence 43, Appl
732	21	75.0	182	4	US-09-489-039A-10417	Sequence 10417, A	805	21	75.0	231	3	US-08-980-357-43	Sequence 43, Appl
733	21	75.0	184	4	US-09-565-423-11	Sequence 11, Appl	806	21	75.0	231	3	US-08-980-357-43	Sequence 4094, Ap
734	21	75.0	185	3	US-08-981-739-133	Sequence 133, App	807	21	75.0	232	3	US-09-134-001C-4094	Sequence 2, Appl
735	21	75.0	185	3	US-08-981-739-136	Sequence 136, App	808	21	75.0	232	3	US-09-094-103-2	Sequence 5010, Ap
736	21	75.0	185	4	US-09-128-026-133	Sequence 133, App	809	21	75.0	232	4	US-09-583-110-5010	Sequence 2, Appl
737	21	75.0	185	4	US-09-128-026-136	Sequence 136, App	810	21	75.0	233	2	US-08-771-783-2	Sequence 177, App
738	21	75.0	185	4	US-09-565-423-17	Sequence 17, Appl	811	21	75.0	236	4	US-08-311-731A-177	Sequence 6873, Ap
739	21	75.0	185	4	US-09-220-616-133	Sequence 133, App	812	21	75.0	236	4	US-09-107-532A-6873	Sequence 4234, Ap
740	21	75.0	185	4	US-09-220-616-136	Sequence 136, App	813	21	75.0	237	4	US-09-443-184-49	Sequence 49, Appl
741	21	75.0	185	4	US-09-220-527-133	Sequence 133, App	814	21	75.0	238	3	US-09-270-767-42453	Sequence 42453, A
742	21	75.0	185	4	US-09-220-527-136	Sequence 136, App	815	21	75.0	239	4	US-09-107-532A-4551	Sequence 4551, Ap
743	21	75.0	185	4	US-09-854-864-11	Sequence 11, Appl	816	21	75.0	240	4	US-09-543-681A-6205	Sequence 6205, Ap
744	21	75.0	189	4	US-09-489-039A-8614	Sequence 8614, Ap	817	21	75.0	241	4	US-09-543-681A-6205	Sequence 4495, A
745	21	75.0	191	3	US-09-134-001C-5345	Sequence 5345, Ap	818	21	75.0	243	4	US-09-270-767-4495	Sequence 3779, Ap
746	21	75.0	192	4	US-09-328-352-7687	Sequence 7687, Ap	819	21	75.0	245	3	US-09-134-001C-3779	Sequence 9, Appl
747	21	75.0	192	4	US-08-311-731A-54	Sequence 54, Appl	820	21	75.0	246	2	US-08-271-354-9	Sequence 9, Appl
748	21	75.0	192	4	US-08-311-731A-65	Sequence 65, Appl	821	21	75.0	246	2	US-08-565-861-9	Sequence 23985, A
749	21	75.0	192	4	US-09-489-039A-12016	Sequence 12016, A	822	21	75.0	246	5	PCT-US94-07658-9	Sequence 7743, Ap
750	21	75.0	193	3	US-09-134-001C-5395	Sequence 5395, Ap	823	21	75.0	246	5	US-09-328-352-7743	Sequence 1686, Ap
751	21	75.0	197	3	US-09-813-918-4	Sequence 4, Appl	824	21	75.0	248	4	US-09-710-279-1686	Sequence 30997, A
752	21	75.0	197	4	US-10-060-311-4	Sequence 4, Appl	825	21	75.0	250	4	US-09-282-991A-30997	Sequence 40, Appl
753	21	75.0	199	4	US-09-489-039A-10661	Sequence 10661, A	826	21	75.0	254	4	US-09-285-219A-40	Sequence 44, Appl
754	21	75.0	200	3	US-09-134-001C-3812	Sequence 3812, Ap	827	21	75.0	257	4	US-09-252-991A-32989	Sequence 32989, A
755	21	75.0	201	1	US-07-929-580B-5	Sequence 5, Appl	828	21	75.0	259	4	US-09-134-000C-4642	Sequence 4642, Ap
756	21	75.0	201	4	US-09-198-452A-548	Sequence 548, App	829	21	75.0	259	4	US-09-198-452A-1063	Sequence 1063, Ap
757	21	75.0	205	4	US-09-710-279-3020	Sequence 3020, Ap	830	21	75.0	263	4		

831	21	75.0	264	4	US-09-328-352-6490	Sequence 6490, Ap	904	1	US-08-125-746-3	Sequence 3, Appli
832	21	75.0	264	4	US-09-886-319A-68	Sequence 68, Appl	905	320	US-08-852-730-9	Sequence 9, Appli
833	21	75.0	265	4	US-09-252-991A-29625	Sequence 29625, A	906	320	US-08-948-276-1	Sequence 1, Appli
834	21	75.0	267	4	US-09-252-991A-26894	Sequence 26894, A	907	320	US-09-248-796A-17061	Sequence 17061, A
835	21	75.0	268	4	US-09-328-352-4581	Sequence 4581, Ap	908	320	5225537-4	Patent No. 5225537
836	21	75.0	268	4	US-09-538-092-735	Sequence 735, App	909	323	US-08-404-445-3	Sequence 3, Appli
837	21	75.0	269	4	US-09-634-238-222	Sequence 232, App	910	323	US-09-543-681A-7304	Sequence 7304, Ap
838	21	75.0	270	4	US-09-230-041-2	Sequence 2, Appli	911	324	US-09-328-352-4636	Sequence 4636, Ap
839	21	75.0	271	1	US-08-276-919-10	Sequence 10, Appl	912	324	US-08-956-171E-5209	Sequence 5209, Ap
840	21	75.0	271	1	US-08-776-088-13	Sequence 13, Appl	913	324	US-09-270-767-42432	Sequence 42432, A
841	21	75.0	271	5	PCT-US95-09145A-13	Sequence 13, Appl	914	324	US-08-781-986A-5209	Sequence 5209, Ap
842	21	75.0	272	4	US-09-107-532A-4656	Sequence 4656, Ap	915	325	US-08-276-919-4	Sequence 4, Appli
843	21	75.0	274	1	US-08-776-088-15	Sequence 15, Appl	916	325	US-08-276-919-13	Sequence 13, Appl
844	21	75.0	274	5	PCT-US95-09145A-15	Sequence 15, Appl	917	325	US-08-776-088-4	Sequence 4, Appli
845	21	75.0	276	4	US-09-328-352-6348	Sequence 6348, Ap	918	325	US-08-776-088-18	Sequence 18, Appl
846	21	75.0	277	4	US-09-270-767-61746	Sequence 61746, A	919	325	US-09-186-276B-21	Sequence 21, Appl
847	21	75.0	279	4	US-09-543-681A-7858	Sequence 7858, Ap	920	325	US-08-842-445-21	Sequence 21, Appl
848	21	75.0	281	1	US-08-487-748A-9	Sequence 9, Appli	921	325	US-09-186-188B-21	Sequence 21, Appl
849	21	75.0	281	1	US-08-487-748A-10	Sequence 10, Appl	922	325	PCT-US95-09145A-4	Sequence 4, Appli
850	21	75.0	281	3	US-08-480-070C-10	Sequence 10, Appl	923	325	PCT-US95-09145A-18	Sequence 18, Appl
851	21	75.0	281	3	US-08-829-525-10	Sequence 10, Appl	924	326	US-09-186-276B-33	Sequence 33, Appl
852	21	75.0	281	3	US-08-609-583A-10	Sequence 10, Appl	925	326	US-08-842-445-33	Sequence 33, Appl
853	21	75.0	281	3	US-08-937-399-10	Sequence 10, Appl	926	326	US-09-186-188B-33	Sequence 33, Appl
854	21	75.0	281	4	US-09-310-367-10	Sequence 10, Appl	927	326	US-09-134-000C-4989	Sequence 4989, Ap
855	21	75.0	281	4	US-09-032-337-10	Sequence 10, Appl	928	327	US-09-324-096A-2	Sequence 2, Appli
856	21	75.0	281	4	US-09-252-991A-20589	Sequence 20589, A	929	327	US-09-324-096A-4	Sequence 4, Appli
857	21	75.0	281	4	US-09-464-231-10	Sequence 10, Appl	930	327	US-09-324-096A-6	Sequence 6, Appli
858	21	75.0	282	4	US-09-561-077C-32	Sequence 32, Appl	931	327	US-09-386-629-8	Sequence 8, Appli
859	21	75.0	282	4	US-09-221-014-32	Sequence 32, Appl	932	329	US-08-887-534A-24	Sequence 24, Appl
860	21	75.0	282	4	US-09-248-796A-17525	Sequence 17525, A	933	329	US-09-527-431-24	Sequence 24, Appl
861	21	75.0	286	3	US-09-134-001C-4602	Sequence 4602, Ap	934	329	US-09-446-861-24	Sequence 24, Appl
862	21	75.0	286	4	US-09-134-000C-4899	Sequence 4899, Ap	935	332	US-09-134-001C-5229	Sequence 5229, Ap
863	21	75.0	287	4	US-09-107-532A-4644	Sequence 4644, Ap	936	333	US-09-252-991A-29336	Sequence 29336, A
864	21	75.0	287	4	US-09-489-039A-9585	Sequence 9585, Ap	937	334	US-09-583-110-3037	Sequence 3037, Ap
865	21	75.0	288	3	US-09-273-839A-8	Sequence 8, Appli	938	336	US-09-252-991A-29624	Sequence 29624, A
866	21	75.0	290	4	US-09-252-991A-32494	Sequence 32494, A	939	336	US-09-583-110-3957	Sequence 3957, Ap
867	21	75.0	293	4	US-09-934-899-2	Sequence 2, Appli	940	337	US-09-252-991A-24654	Sequence 24654, A
868	21	75.0	293	4	US-09-934-868-22	Sequence 22, Appl	941	338	US-09-935-263-8	Sequence 8, Appli
869	21	75.0	293	4	US-09-270-767-46182	Sequence 46182, A	942	338	US-09-594-185-8	Sequence 8, Appli
870	21	75.0	293	4	US-09-248-796A-18524	Sequence 18524, A	943	338	US-09-107-532A-6222	Sequence 6222, Ap
871	21	75.0	294	4	US-09-252-991A-26953	Sequence 26953, A	944	338	US-10-033-078-8	Sequence 8, Appli
872	21	75.0	296	4	US-09-583-110-2963	Sequence 2963, Ap	945	339	US-09-328-352-4674	Sequence 4674, Ap
873	21	75.0	297	4	US-09-107-532A-5306	Sequence 5306, Ap	946	339	US-09-107-532A-5514	Sequence 5514, Ap
874	21	75.0	297	4	US-09-134-000C-6433	Sequence 6433, Ap	947	340	US-09-107-532A-6283	Sequence 6283, Ap
875	21	75.0	299	4	US-09-543-681A-4901	Sequence 4901, Ap	948	341	US-09-134-000C-4747	Sequence 4747, Ap
876	21	75.0	300	4	US-09-252-991A-28063	Sequence 28063, A	949	341	US-09-134-000C-6285	Sequence 6285, Ap
877	21	75.0	301	4	US-09-583-110-3523	Sequence 3523, Ap	950	341	US-09-270-767-44982	Sequence 44982, A
878	21	75.0	303	2	US-08-599-171A-32	Sequence 32, Appl	951	342	US-09-489-039A-13595	Sequence 13595, A
879	21	75.0	303	2	US-08-646-590B-32	Sequence 32, Appl	952	342	US-09-540-236-2385	Sequence 2385, Ap
880	21	75.0	303	3	US-09-069-226-32	Sequence 32, Appl	953	343	US-09-489-039A-9368	Sequence 9368, Ap
881	21	75.0	303	3	US-09-412-184-32	Sequence 32, Appl	954	344	US-09-199-637A-108	Sequence 108, App
882	21	75.0	303	4	US-09-107-532A-5192	Sequence 5192, Ap	955	344	US-09-540-236-2452	Sequence 2452, Ap
883	21	75.0	304	3	US-09-632-947B-7	Sequence 7, Appli	956	346	US-08-774-169-3	Sequence 3, Appli
884	21	75.0	304	4	US-09-489-039A-10303	Sequence 10303, A	957	348	US-08-855-910-8	Sequence 8, Appli
885	21	75.0	310	4	US-09-372-934-4	Sequence 4, Appli	958	349	US-09-489-039A-13939	Sequence 13939, A
886	21	75.0	310	4	US-09-766-916-4	Sequence 4, Appli	959	352	US-09-107-532A-4936	Sequence 4936, Ap
887	21	75.0	310	4	US-09-766-898-4	Sequence 4, Appli	960	352	US-09-710-279-216	Sequence 216, App
888	21	75.0	310	4	US-09-637-040C-4	Sequence 4, Appli	961	353	US-09-252-991A-32769	Sequence 32769, A
889	21	75.0	311	4	US-09-252-991A-19489	Sequence 19489, A	962	354	US-08-276-919-2	Sequence 2, Appli
890	21	75.0	312	4	US-09-540-236-2460	Sequence 2460, Ap	963	354	US-07-868-353A-13	Sequence 13, Appl
891	21	75.0	313	3	US-09-134-001C-5185	Sequence 5185, Ap	964	354	US-08-776-088-2	Sequence 2, Appli
892	21	75.0	314	4	US-09-248-796A-18201	Sequence 18201, A	965	354	US-08-407-804-2	Sequence 6, Appli
893	21	75.0	315	1	US-07-729-099-1	Sequence 1, Appli	966	354	US-09-134-807-32	Sequence 22, Appl
894	21	75.0	315	3	US-08-257-392-1	Sequence 1, Appli	967	354	US-09-325-320-2	Sequence 22, Appl
895	21	75.0	315	3	US-08-770-035-1	Sequence 1, Appli	968	354	US-09-585-109-2	Sequence 2, Appli
896	21	75.0	316	3	US-09-078-691-2	Sequence 2, Appli	969	354	US-08-976-063E-4	Sequence 2, Appli
897	21	75.0	316	4	US-09-387-375-9	Sequence 9, Appli	970	354	PCT-US95-09145A-2	Sequence 4, Appli
898	21	75.0	316	4	US-10-041-400A-9	Sequence 9, Appli	971	354	PCT-US95-09145A-6	Sequence 6, Appli
899	21	75.0	318	4	US-08-948-276-3	Sequence 3, Appli	972	355	US-09-252-991A-33034	Sequence 33034, A
900	21	75.0	319	1	US-08-125-746-1	Sequence 1, Appli	973	355	US-09-270-767-42551	Sequence 42551, A
901	21	75.0	319	2	US-08-149-975A-2	Sequence 2, Appli	974	355	US-08-960-780-19	Sequence 19, Appl
902	21	75.0	319	4	US-08-948-276-2	Sequence 2, Appli	975	357	US-09-073-898-19	Sequence 19, Appl
903	21	75.0	319	4	US-08-311-731A-63	Sequence 63, Appl	976	357		

977 21 75.0 357 4 US-09-850-351A-19 Sequence 19, Appl  
 978 21 75.0 359 4 US-09-538-092-111 Sequence 111, App  
 979 21 75.0 360 4 US-09-252-991A-20193 Sequence 20193, A  
 980 21 75.0 361 3 US-09-120-365-75 Sequence 75, Appl  
 981 21 75.0 361 3 US-09-515-039-75 Sequence 75, Appl  
 982 21 75.0 361 4 US-09-724-566A-75 Sequence 75, Appl  
 983 21 75.0 361 4 US-09-583-110-5129 Sequence 5129, Ap  
 984 21 75.0 362 3 US-09-625-188-8 Sequence 8, Appl  
 985 21 75.0 362 4 US-09-442-349A-109 Sequence 109, App  
 986 21 75.0 364 4 US-09-710-279-3076 Sequence 3076, Ap  
 987 21 75.0 365 4 US-09-442-349A-108 Sequence 108, App  
 988 21 75.0 367 4 US-09-134-000C-6173 Sequence 6173, Ap  
 989 21 75.0 369 1 US-08-232-238A-2 Sequence 2, Appl  
 990 21 75.0 369 2 US-08-468-865-2 Sequence 2, Appl  
 991 21 75.0 369 2 US-08-411-043-2 Sequence 2, Appl  
 992 21 75.0 373 3 US-09-134-001C-4798 Sequence 4798, Ap  
 993 21 75.0 373 4 US-09-248-796A-18227 Sequence 18227, A  
 994 21 75.0 374 3 US-08-878-801-2 Sequence 2, Appl  
 995 21 75.0 374 4 US-09-442-349A-1 Sequence 1, Appl  
 996 21 75.0 374 4 US-09-442-349A-2 Sequence 2, Appl  
 997 21 75.0 374 4 US-09-442-349A-3 Sequence 3, Appl  
 998 21 75.0 374 4 US-09-442-349A-4 Sequence 4, Appl  
 999 21 75.0 374 4 US-09-442-349A-5 Sequence 5, Appl  
 1000 21 75.0 374 4 US-09-442-349A-6 Sequence 6, Appl

ALIGNMENTS

RESULT 1  
 US-08-972-902-7  
 ; Sequence 7, Application US/08972902  
 ; Patent No. 6099848  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Frankel, Fred R. A.  
 ; APPLICANT: Portnoy, Daniel A.  
 ; TITLE OF INVENTION: BACTERIAL VACCINE VECTOR AND METHODS OF  
 ; TITLE OF INVENTION: USE THEREOF  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
 ; STREET: One Commerce Square, 2005 Market Street, 22nd  
 ; CITY: Philadelphia  
 ; STATE: PA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 19103-7086  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/972,902  
 ; FILING DATE: 18-NOV-1997  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Doyle-Leary, Kathryn  
 ; REGISTRATION NUMBER: 36,317  
 ; REFERENCE/DOCKET NUMBER: 9596-51  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-965-1284  
 ; TELEFAX: 215-567-2991  
 ; TELEX: 831-494  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 282 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-972-902-7

Query Match 96.4%; Score 27; DB 3; Length 282;  
 Best Local Similarity 83.3%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 GYXVEE 6  
 Db 64 GYTVEE 69  
 RESULT 2  
 US-09-520-207-7  
 ; Sequence 7, Application US/09520207  
 ; Patent No. 6504020  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FRANKEL, Fred R.  
 ; APPLICANT: PORTNOY, Daniel A.  
 ; TITLE OF INVENTION: A BACTERIAL VACCINE VECTOR AND METHODS OF USE THEREOF  
 ; FILE REFERENCE: 9596-51U1 (209596.0337)  
 ; CURRENT APPLICATION NUMBER: US/09/520,207  
 ; CURRENT FILING DATE: 2000-03-07  
 ; PRIOR APPLICATION NUMBER: U.S. 08/972,902  
 ; PRIOR FILING DATE: 1997-11-18  
 ; PRIOR APPLICATION NUMBER: PCT/US98/24357  
 ; PRIOR FILING DATE: 1998-11-13  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 7  
 ; LENGTH: 282  
 ; TYPE: PRT  
 ; ORGANISM: Listeria monocytogenes  
 ; US-09-520-207-7  
 Query Match 96.4%; Score 27; DB 4; Length 282;  
 Best Local Similarity 83.3%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 GYXVEE 6  
 Db 64 GYTVEE 69  
 RESULT 3  
 US-10-136-253-7  
 ; Sequence 7, Application US/10136253  
 ; Patent No. 6635749  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FRANKEL, Fred R.  
 ; APPLICANT: PORTNOY, Daniel A.  
 ; TITLE OF INVENTION: ISOLATED NUCLEIC ACIDS COMPRISING LISTERIA DAL AND DAT GENES  
 ; FILE REFERENCE: 053893-5011-02  
 ; CURRENT APPLICATION NUMBER: US/10/136,253  
 ; CURRENT FILING DATE: 2002-05-01  
 ; PRIOR APPLICATION NUMBER: 09/520,207  
 ; PRIOR FILING DATE: 2000-03-07  
 ; PRIOR APPLICATION NUMBER: PCT/US98/24357  
 ; PRIOR FILING DATE: 1998-11-13  
 ; PRIOR APPLICATION NUMBER: 08/972,902  
 ; PRIOR FILING DATE: 1997-11-18  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 7  
 ; LENGTH: 282  
 ; TYPE: PRT  
 ; ORGANISM: Listeria monocytogenes  
 ; US-10-136-253-7  
 Query Match 96.4%; Score 27; DB 4; Length 282;  
 Best Local Similarity 83.3%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 GYXVEE 6  
 Db 64 GYTVEE 69





; LOCATION: (601)...(601)  
; OTHER INFORMATION: Xaa at position 601 is alanine or proline  
US-09-391-340-10

Query Match 96.4%; Score 27; DB 4; Length 762;  
Best Local Similarity 83.3%; Pred. No. 2,5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
Db 689 GYSVEE 694

RESULT 8  
US-08-313-050-18  
; Sequence 18, Application US/08313050  
; Patent No. 5585256  
; GENERAL INFORMATION:  
; APPLICANT: Dorreich, Kurt  
; APPLICANT: Christensen, Flemming M.  
; APPLICANT: Schnell, Yvette  
; APPLICANT: Mischler, Marcel  
; APPLICANT: Dalboge, Henrik  
; APPLICANT: Helldt-Hansen, Hans P.  
; TITLE OF INVENTION: A NOVEL ENZYME AND A DNA SEQUENCE  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 55852560 No. 5585256disk of No. 5585256th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/313.050  
; FILING DATE: 05-OCT-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DK 0420/92  
; FILING DATE: 27-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DK PCT/DK93/00109  
; FILING DATE: 29-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lowney Dr., Karen A.  
; REGISTRATION NUMBER: 31,274  
; REFERENCE/DOCKET NUMBER: 3730.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 53 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-313-050-18

Query Match 92.9%; Score 26; DB 1; Length 53;  
Best Local Similarity 83.3%; Pred. No. 24;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
Db 41 GYVVEE 46

RESULT 9

US-08-808-982-8  
; Sequence 8, Application US/08808982  
; Patent No. 5939271  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: Leonardo, E. David  
; APPLICANT: Hink, Lindsay  
; APPLICANT: Masu, Masayuki  
; APPLICANT: Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,982  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-808-982-8

Query Match 92.9%; Score 26; DB 2; Length 102;  
Best Local Similarity 83.3%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
Db 84 GYLVEE 89

RESULT 10  
US-09-306-902A-8  
; Sequence 8, Application US/09306902A  
; Patent No. 6277585  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: Leonardo, E. David  
; APPLICANT: Hink, Lindsay  
; APPLICANT: Masu, Masayuki  
; APPLICANT: Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/306,902A  
FILING DATE: 07-May-1999  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UC96-217  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-306-902A-8

Query Match 92.9%; Score 26; DB 3; Length 102;  
Best Local Similarity 83.3%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 84 GYLVEE 89

## RESULT 11

US-09-107-532A-5829  
Sequence 5829, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION/  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5829:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 107 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: Protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...107  
SEQUENCE DESCRIPTION: SEQ ID NO: 5829:  
US-09-107-532A-5829

Query Match 92.9%; Score 26; DB 4; Length 107;  
Best Local Similarity 66.7%; Pred. No. 52;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 86 GYSIEE 91

## RESULT 12

US-09-173-300-13  
Sequence 13, Application US/09173300  
Patent No. 6451581  
GENERAL INFORMATION:  
APPLICANT: Falco, Saverio Carl  
APPLICANT: Hitz, William D.  
APPLICANT: Kinney, Anthony J.  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Rafalski, J. Antoni  
TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC  
ENZYMES  
FILE REFERENCE: BB-1126  
CURRENT APPLICATION NUMBER: US/09/173,300  
CURRENT FILING DATE: 1998-10-15  
EARLIER APPLICATION NUMBER: 60/063,423  
EARLIER FILING DATE: 1997 October 28  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: Microsoft Word Version 7.0A  
SEQ ID NO 13  
LENGTH: 115  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (104)  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (112)  
US-09-173-300-13

Query Match 92.9%; Score 26; DB 4; Length 115;  
Best Local Similarity 83.3%; Pred. No. 56;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 33 GYKVEE 38

## RESULT 13

US-09-585-173B-20  
Sequence 20, Application US/09585173B  
Patent No. 6570063  
GENERAL INFORMATION:  
APPLICANT: Butler, Karlene  
APPLICANT: Famodu, Omolayo O.  
APPLICANT: Gutteridge, Steven  
APPLICANT: Maxwell, Carl  
TITLE OF INVENTION: Magnesium Chelataase  
FILE REFERENCE: BB1370 US NA  
CURRENT APPLICATION NUMBER: US/09/585,173B  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/137,461

```
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Zea mays
US-09-585-173B-20

Query Match          92.9%; Score 26; DB 4; Length 118;
Best Local Similarity 83.3%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      23 GYLVEE 28

RESULT 14
US-09-134-000C-3570
; Sequence 3570, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3570
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3570

Query Match          92.9%; Score 26; DB 4; Length 128;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      92 GYRVEE 97

RESULT 15
US-09-107-532A-6268
; Sequence 6268, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/107,532A
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3570
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3570

Query Match          92.9%; Score 26; DB 4; Length 128;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      92 GYRVEE 97

RESULT 16
US-09-585-173B-22
; Sequence 22, Application US/09585173B
; Patent No. 6570063
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Gutteridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: BB1370 US NA
; CURRENT APPLICATION NUMBER: US/09/585,173B
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Zea mays
US-09-585-173B-22

Query Match          92.9%; Score 26; DB 4; Length 260;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      26 GYLVEE 31

RESULT 17
US-09-173-300-15
; Sequence 15, Application US/09173300
; Patent No. 6451581
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
```

```
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6268:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...238
; SEQUENCE DESCRIPTION: SEQ ID NO: 6268:
US-09-107-532A-6268

Query Match          92.9%; Score 26; DB 4; Length 238;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      168 GYVEEE 173

RESULT 16
US-09-585-173B-22
; Sequence 22, Application US/09585173B
; Patent No. 6570063
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Gutteridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: BB1370 US NA
; CURRENT APPLICATION NUMBER: US/09/585,173B
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Zea mays
US-09-585-173B-22

Query Match          92.9%; Score 26; DB 4; Length 260;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      26 GYLVEE 31

RESULT 17
US-09-173-300-15
; Sequence 15, Application US/09173300
; Patent No. 6451581
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
```

```
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/09/173,300
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 60/063,423
; EARLIER FILING DATE: 1997 October 28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 15
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-173-300-15

Query Match          92.9%; Score 26; DB 4; Length 307;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      230 GYQVEE 235

RESULT 18
US-09-564-805-231
; Sequence 231, Application US/09564805
; Patent No. 6333403.
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 231
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Synechocystis sp.
US-09-564-805-231

Query Match          92.9%; Score 26; DB 3; Length 326;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      146 GYRVEE 151

RESULT 19
US-08-183-214-12
; Sequence 12, Application US/08183214
; Patent No. 5716816
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Stanley, Sally J.
; APPLICANT: Nightingale, Maria S.
; APPLICANT: Murtagh, Jr., James J.
; APPLICANT: Monaco, Lucia
```

```
; APPLICANT: Takada, Tatsuyuki
; TITLE OF INVENTION: CLONES ENCODING MAMMALIAN
; TITLE OF INVENTION: ADP-RIBOSYLARGININE HYDROLASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,214
; FILING DATE: 14-JAN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,231
; FILING DATE: 22-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-183-214-12

Query Match          92.9%; Score 26; DB 1; Length 345;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      198 GYFVEE 203

RESULT 20
US-09-134-001C-4405
; Sequence 4405, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4405
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4405

Query Match          92.9%; Score 26; DB 3; Length 357;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 GYXVEE 6
      |||||
Db     321 GYVEE 326

RESULT 21
US-09-710-279-18
; Sequence 18, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-18

Query Match      92.9%; Score 26; DB 4; Length 358;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      |||||
Db     272 GYVEE 277

RESULT 22
US-09-134-001C-5403
; Sequence 5403, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5403
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5403

Query Match      92.9%; Score 26; DB 3; Length 362;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      |||||
Db     276 GYVEE 281

RESULT 23
US-09-543-681A-7314
; Sequence 7314, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7314
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7314

Query Match      92.9%; Score 26; DB 4; Length 369;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      |||||
Db     291 GYIEE 296

RESULT 24
US-09-270-767-59898
; Sequence 59898, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59898
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-59898

Query Match      92.9%; Score 26; DB 4; Length 378;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      |||||
Db     334 GYIVEE 339

RESULT 25
US-09-673-395A-363
; Sequence 363, Application US/09673395A
; Patent No. 6620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKI, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 363
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-395A-363
```

```
Query Match          92.9%; Score 26; DB 4; Length 381;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      97 GYGVVE 102

RESULT 26
US-09-543-681A-7895
; Sequence 7895, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7895
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7895

Query Match          92.9%; Score 26; DB 4; Length 412;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      228 GYAIIE 233

RESULT 27
US-09-485-529-6
; Sequence 6, Application US/09485529
; Patent No. 6762348
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Richards, Donald E
; APPLICANT: Peng, Jinrong
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development
; FILE REFERENCE: 620-91
; CURRENT APPLICATION NUMBER: US/09/485,529
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/GB98/02383
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: GB 971192.0
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-485-529-6

Query Match          92.9%; Score 26; DB 4; Length 425;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      395 GYKVEE 400

RESULT 28
```

```
US-09-134-001C-5322
; Sequence 5322, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5322
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5322

Query Match          92.9%; Score 26; DB 3; Length 460;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      24 GYSIEE 29

RESULT 29
US-09-117-853-2
; Sequence 2, Application US/09117853
; Patent No. 6307126
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; CURRENT FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-117-853-2

Query Match          92.9%; Score 26; DB 3; Length 532;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      502 GYRVEE 507

RESULT 30
US-09-911-154-2
; Sequence 2, Application US/09911154
; Patent No. 6478809
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
```

```
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,154
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 532
; ORGANISM: Arabidopsis thaliana
US-09-911-154-2

Query Match          92.9%; Score 26; DB 4; Length 532;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      502 GYRVEE 507

RESULT 31
US-09-485-529-2
; Sequence 2, Application US/09485529
; Patent No. 6762348
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Richards, Donald E
; APPLICANT: Peng, Jinrong
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development
; FILE REFERENCE: 620-91
; CURRENT APPLICATION NUMBER: US/09/485,529
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/GB98/02383
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: GB 9717192.0
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-485-529-2

Query Match          92.9%; Score 26; DB 4; Length 532;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      502 GYRVEE 507

RESULT 32
US-09-911-514-2
; Sequence 2, Application US/09911514
; Patent No. 6794560
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,514
; CURRENT FILING DATE: 2001-07-25
;
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-158
; CURRENT APPLICATION NUMBER: US/09/911,514
; CURRENT FILING DATE: 2001-07-25
;
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development
; FILE REFERENCE: 620-91
; CURRENT APPLICATION NUMBER: US/09/485,529
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/GB98/02383
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: GB 9717192.0
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-485-529-7

Query Match          92.9%; Score 26; DB 4; Length 623;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      593 GYRVEE 598

RESULT 34
US-09-485-529-8
; Sequence 8, Application US/09485529
; Patent No. 6762348
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Richards, Donald E
; APPLICANT: Peng, Jinrong
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development
; FILE REFERENCE: 620-91
; CURRENT APPLICATION NUMBER: US/09/485,529
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/GB98/02383
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: GB 9717192.0
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 108
```

```
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-911-514-2

Query Match          92.9%; Score 26; DB 4; Length 532;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      502 GYRVEE 507

RESULT 33
US-09-485-529-7
; Sequence 7, Application US/09485529
; Patent No. 6762348
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Richards, Donald E
; APPLICANT: Peng, Jinrong
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development
; FILE REFERENCE: 620-91
; CURRENT APPLICATION NUMBER: US/09/485,529
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/GB98/02383
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: GB 9717192.0
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-485-529-7

Query Match          92.9%; Score 26; DB 4; Length 623;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      593 GYRVEE 598

RESULT 34
US-09-485-529-8
; Sequence 8, Application US/09485529
; Patent No. 6762348
; GENERAL INFORMATION:
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Richards, Donald E
; APPLICANT: Peng, Jinrong
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development
; FILE REFERENCE: 620-91
; CURRENT APPLICATION NUMBER: US/09/485,529
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/GB98/02383
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: GB 9717192.0
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 108
```



; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 630

; TYPE: PRT

; ORGANISM: Zea mays

US-09-485-529-8

Query Match 92.9%; Score 26; DB 4; Length 630;

Best Local Similarity 83.3%; Pred. No. 3.5e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

|||||

Db 597 GYRVEE 602

RESULT 35

US-09-489-847-200

; Sequence 200, Application US/09489847

; Patent No. 6476195

; GENERAL INFORMATION:

; APPLICANT: Rosen et al

; TITLE OF INVENTION: 98 Human Secreted Proteins

; FILE REFERENCE: P2031P1

; CURRENT APPLICATION NUMBER: US/09/489,847

; CURRENT FILING DATE: 2000-01-24

; EARLIER APPLICATION NUMBER: PCT/US99/17130

; EARLIER FILING DATE: 1999-07-29

; EARLIER APPLICATION NUMBER: 60/094,657

; EARLIER FILING DATE: 1998-07-30

; EARLIER APPLICATION NUMBER: 60/095,486

; EARLIER FILING DATE: 1998-08-05

; EARLIER APPLICATION NUMBER: 60/096,319

; EARLIER FILING DATE: 1998-08-12

; EARLIER APPLICATION NUMBER: 60/095,454

; EARLIER FILING DATE: 1998-08-06

; EARLIER APPLICATION NUMBER: 60/095,455

; EARLIER FILING DATE: 1998-08-06

; NUMBER OF SEQ ID NOS: 376

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 200

; LENGTH: 698

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-489-847-200

Query Match

Best Local Similarity 92.9%; Score 26; DB 4; Length 698;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

|||||

Db 321 GYVEE 326

RESULT 36

US-09-270-767-44461

; Sequence 44461, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 44461

; LENGTH: 705

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-44461

Query Match

Best Local Similarity 92.9%; Score 26; DB 4; Length 705;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

|||||

Db 334 GYVEE 339

RESULT 37

US-09-252-991A-24946

; Sequence 24946, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 24946

; LENGTH: 710

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24946

Query Match

Best Local Similarity 92.9%; Score 26; DB 4; Length 710;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

|||||

Db 175 GYLVEE 180

RESULT 38

US-08-836-943-2

; Sequence 2, Application US/08836943

; Patent No. 5965391

; GENERAL INFORMATION:

; APPLICANT: Reinscheid, Dieter

; APPLICANT: Eikmanns, Bernhard

; APPLICANT: Sahm, Hermann

; TITLE OF INVENTION: DNA WHICH REGULATES GENE EXPRESSION IN

; TITLE OF INVENTION: CORYNEFORM BACTERIA

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: The Firm of Karl F. Ross, PC

; STREET: 5676 Riverdale Ave.

; CITY: Bronx

; STATE: New York

; COUNTRY: USA

; ZIP: 10471

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/836,943

; FILING DATE: 08-MAY-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Myers, Jonathan

; REGISTRATION NUMBER: 26,963

; REFERENCE/DOCKET NUMBER: 20357

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (718) 884-6600  
 TELEFAX: 718/601-1099  
 TELEX: 620428  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 739 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-836-943-2

Query Match 92.9%; Score 26; DB 2; Length 739;  
 Best Local Similarity 83.3%; Pred. No. 4.2e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 Db 106 GYLVEE 111

RESULT 39  
 US-08-149-103-3  
 ; Sequence 3, Application US/08149103  
 ; Patent No. 5750367  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lawrence C. B. Chan  
 ; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW  
 ; TITLE OF INVENTION: DENSITY LIPOPROTEIN RECEPTORS  
 ; TITLE OF INVENTION: AND METHODS FOR USE OF SUCH  
 ; TITLE OF INVENTION: RECEPTORS  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LYON & LYON  
 ; STREET: 611 West Sixth Street  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90017  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)  
 ; SOFTWARE: WordPerfect (Version 5.1)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/149,103  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA: including application  
 ; PRIOR APPLICATION DATA: described below:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Warburg, Richard J.  
 ; REGISTRATION NUMBER: 32,327  
 ; REFERENCE/DOCKET NUMBER: 204/052  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 846 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; none

Query Match 92.9%; Score 26; DB 1; Length 846;  
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 Db 712 GYNVEE 717

RESULT 40  
 US-08-451-883-3  
 ; Sequence 3, Application US/08451883  
 ; Patent No. 5798209  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lawrence C.B. Chan  
 ; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW DENSITY  
 ; TITLE OF INVENTION: LIPOPROTEIN RECEPTORS AND METHODS FOR  
 ; TITLE OF INVENTION: USE OF SUCH RECEPTORS  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LYON & LYON  
 ; STREET: 633 West Fifth Street, Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071-2066  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: IBM MS-DOS (Version 6.22)  
 ; SOFTWARE: WordPerfect (Version 5.1)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/451,883  
 ; FILING DATE: May 26, 1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; PRIOR APPLICATION DATA: including application  
 ; PRIOR APPLICATION DATA: described below:  
 ; APPLICATION NUMBER: 08/149,103  
 ; FILING DATE: No. 5798209ember 8, 1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Knight, Matthew W.  
 ; REGISTRATION NUMBER: 36,846  
 ; REFERENCE/DOCKET NUMBER: 212/268  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 846 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-451-883-3

Query Match 92.9%; Score 26; DB 1; Length 846;  
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
 Db 712 GYNVEE 717

RESULT 41  
 US-09-107-532A-6829  
 ; Sequence 6829, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street

```
; ; CITY: Waltham
; ; STATE: Massachusetts
; ; COUNTRY: USA
; ; ZIP: 02354
; ;
; ; COMPUTER READABLE FORM:
; ; MEDIUM TYPE: CD/ROM ISO9660
; ; COMPUTER: PC
; ; OPERATING SYSTEM: <Unknown>
; ; SOFTWARE: ASCII
; ;
; ; CURRENT APPLICATION DATA:
; ; FILING APPLICATION NUMBER: US/09/107,532A
; ; FILING DATE: 30-Jun-1998
; ;
; ; PRIOR APPLICATION DATA:
; ; APPLICATION NUMBER: 60/085,598
; ; FILING DATE: 14 May 1998
; ; APPLICATION NUMBER: 60/051571
; ; FILING DATE: July 2, 1997
; ;
; ; ATTORNEY/AGENT INFORMATION:
; ; NAME: Ariniello, Pamela Deneke
; ; REGISTRATION NUMBER: 40,489
; ; REFERENCE/DOCKET NUMBER: GTC-012
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (781)893-5007
; ; TELEFAX: (781)893-8277
; ;
; ; INFORMATION FOR SEQ ID NO: 6829:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 859 amino acids
; ; TYPE: amino acid
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: protein
; ; HYPOTHETICAL: YES
; ; ORIGINAL SOURCE:
; ; ORGANISM: Enterococcus faecium
; ; FEATURE:
; ; NAME/KEY: misc feature
; ; LOCATION: (B) LOCATION 1...859
; ; SEQUENCE DESCRIPTION: SEQ ID NO: 6829:
; ;
; ; US-09-107-532A-6829
; ;
; ; Query Match 92.9%; Score 26; DB 4; Length 859;
; ; Best Local Similarity 66.7%; Pred. No. 4.9e+02;
; ; Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
; ;
; ; QY 1 GYXVEE 6
; ; ||||
; ; Db 240 GYTIEE 245
; ;
; ; RESULT 42
; ; US-09-543-681A-4814
; ; Sequence 4814, Application US/09543681A
; ; Patent No. 6605709
; ; GENERAL INFORMATION:
; ; APPLICANT: GARY BRETON
; ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; ; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; ; FILE REFERENCE: 2709.1002-001
; ; CURRENT APPLICATION NUMBER: US/09/543,681A
; ; CURRENT FILING DATE: 2000-04-05
; ; PRIOR APPLICATION NUMBER: US 60/128,706
; ; PRIOR FILING DATE: 1999-04-09
; ; NUMBER OF SEQ ID NOS: 8344
; ; SEQ ID NO 4814
; ; LENGTH: 869
; ; TYPE: PRP
; ; ORGANISM: Proteus mirabilis
; ;
; ; US-09-543-681A-4814
; ;
; ; Query Match 92.9%; Score 26; DB 4; Length 869;
; ; Best Local Similarity 83.3%; Pred. No. 5e+02;
; ; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
; ;
; ; QY 1 GYXVEE 6
; ;
; ; Db 739 GYNVEE 744
; ;
; ; RESULT 44
; ; US-08-894-489-2
; ; Sequence 2, Application US/08894489
; ; Patent No. 6174527
; ; GENERAL INFORMATION:
; ; APPLICANT: Wilson, James M.
; ; APPLICANT: Kozarsky, Karen F.
; ; APPLICANT: Strauss, Jerome F.
; ; TITLE OF INVENTION: Methods and Compositions for Gene
; ; TITLE OF INVENTION: Therapy for the Treatment of Defects in Lipoprotein
; ; NUMBER OF SEQUENCES: 8
; ; CORRESPONDENCE ADDRESS:
; ; ADDRESSEE: Howson and Howson
; ; STREET: Spring House Corporate Cntr., PO Box 457
; ; CITY: Spring House
; ; STATE: Pennsylvania
; ; COUNTRY: USA
; ; ZIP: 19477
; ;
; ; COMPUTER READABLE FORM:
; ; MEDIUM TYPE: Floppy disk
; ; COMPUTER: IBM PC compatible
; ; OPERATING SYSTEM: PC-DOS/MS-DOS
; ; SOFTWARE: PatentIn Release #1.0, Version #1.30
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/08/393,734
; ; FILING DATE:
; ; CLASSIFICATION: 424
; ; ATTORNEY/AGENT INFORMATION:
; ; NAME: Bak, Mary E.
; ; REGISTRATION NUMBER: 31,215
; ; REFERENCE/DOCKET NUMBER: UPNH1254USA
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: 215-540-9200
; ; TELEFAX: 215-540-5818
; ; INFORMATION FOR SEQ ID NO: 2:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 873 amino acids
; ; TYPE: amino acid
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: protein
; ;
; ; US-08-393-734-2
; ;
; ; Query Match 92.9%; Score 26; DB 1; Length 873;
; ; Best Local Similarity 83.3%; Pred. No. 5e+02;
; ; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
; ;
; ; QY 1 GYXVEE 6
; ; ||||
; ; Db 739 GYNVEE 744
; ;
; ; RESULT 44
; ; US-08-894-489-2
; ; Sequence 2, Application US/08894489
; ; Patent No. 6174527
; ; GENERAL INFORMATION:
; ; APPLICANT: Wilson, James M.
; ; APPLICANT: Kozarsky, Karen F.
; ; APPLICANT: Strauss, Jerome F.
; ; TITLE OF INVENTION: Methods and Compositions for Gene
; ; TITLE OF INVENTION: Therapy for the Treatment of Defects in Lipoprotein
; ; NUMBER OF SEQUENCES: 8
; ; CORRESPONDENCE ADDRESS:
; ; ADDRESSEE: Howson and Howson
; ; STREET: Spring House Corporate Cntr., PO Box 457
; ; CITY: Spring House
; ; STATE: Pennsylvania
```

; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,489  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/393,734  
; FILING DATE: 24-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: GNVN.009CIP1USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 873 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-894-489-2

Query Match 92.9%; Score 26; DB 3; Length 873;  
Best Local Similarity 83.3%; Pred. No. 5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 739 GYNVEE 744

RESULT 45  
US-08-808-982-7  
; Sequence 7, Application US/08808982  
; Patent No. 5939271  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: Leonardo, E. David  
; APPLICANT: Hink, Lindsey  
; APPLICANT: Masu, Masayuki  
; APPLICANT: Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,982  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 943 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; US-08-808-982-7  
  
Query Match 92.9%; Score 26; DB 2; Length 943;  
Best Local Similarity 83.3%; Pred. No. 5.4e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GYXVEE 6  
|||  
Db 734 GYLVEE 739  
  
RESULT 46  
US-09-306-902A-7  
; Sequence 7, Application US/09306902A  
; Patent No. 6277585  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: Leonardo, E. David  
; APPLICANT: Hink, Lindsey  
; APPLICANT: Masu, Masayuki  
; APPLICANT: Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/306,902A  
; FILING DATE: 07-May-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 943 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-306-902A-7

Query Match 92.9%; Score 26; DB 3; Length 943;  
Best Local Similarity 83.3%; Pred. No. 5.4e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 734 GYLVEE 739

RESULT 47

## US-10-140-002-146

; Sequence 146, Application US/10140002  
; Patent No. 6725730  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C59

; CURRENT APPLICATION NUMBER: US/10/140,002

; CURRENT FILING DATE: 2002-05-06

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 146

; LENGTH: 945

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-140-002-146

Query Match 92.9%; Score 26; DB 4; Length 945;

Best Local Similarity 83.3%; Pred. No. 5.4e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

Db 736 GYLVEE 741

## RESULT 48

US-09-134-000C-6090

; Sequence 6090, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 032796-032

; CURRENT APPLICATION NUMBER: US/09/134,000C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 6812

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6090

; LENGTH: 1074

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-134-000C-6090

Query Match 92.9%; Score 26; DB 4; Length 1074;

Best Local Similarity 66.7%; Pred. No. 6.3e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

Db 456 GYSIEE 461

## RESULT 49

US-09-107-532A-5979

; Sequence 5979, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; FILE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD/ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 5979:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 174 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...174

; SEQUENCE DESCRIPTION: SEQ ID NO: 5979:

US-09-107-532A-5979

Query Match 89.3%; Score 25; DB 4; Length 174;

Best Local Similarity 83.3%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

Db 33 GYPVEE 38

## RESULT 50

US-09-134-000C-3409

; Sequence 3409, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 032796-032

; CURRENT APPLICATION NUMBER: US/09/134,000C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15

```
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3409
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3409

Query Match      89.3%; Score 25; DB 4; Length 178;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
Db      129 GYRIEE 134

RESULT 51
US-09-583-110-3479
; Sequence 3479, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3479
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3479

Query Match      89.3%; Score 25; DB 4; Length 240;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
Db      23 GYLIIE 28

RESULT 52
US-09-564-805-230
; Sequence 230, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 230
; LENGTH: 311
```

```
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-564-805-230

Query Match      89.3%; Score 25; DB 3; Length 311;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
Db      153 GYRIEE 158

RESULT 53
US-09-393-858-20
; Sequence 20, Application US/09393858
; Patent No. 6627747
; GENERAL INFORMATION:
; APPLICANT: Fritz, Christian
; APPLICANT: Youngman, Philip
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
; FILE REFERENCE: 06286-088001
; CURRENT APPLICATION NUMBER: US/09/393,858
; CURRENT FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: 60/099,578
; PRIOR FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-393-858-20

Query Match      89.3%; Score 25; DB 4; Length 311;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
Db      153 GYRIEE 158

RESULT 54
US-10-190-279-20
; Sequence 20, Application US/10190279
; Patent No. 6749858
; GENERAL INFORMATION:
; APPLICANT: Fritz, Christian
; APPLICANT: Youngman, Philip
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
; FILE REFERENCE: 06286-088001
; CURRENT APPLICATION NUMBER: US/10/190,279
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US/09/393,858
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: 60/099,578
; PRIOR FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-190-279-20

Query Match      89.3%; Score 25; DB 4; Length 311;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
```

Db 153 GYRIE 158

RESULT 55

US-09-252-991A-26418

; Sequence 26418, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 26418

; LENGTH: 346

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26418

Query Match 89.3%; Score 25; DB 4; Length 346;

Best Local Similarity 66.7%; Pred. No. 3.2e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6

Db 130 GYQIEE 135

RESULT 56

US-09-489-039A-10041

; Sequence 10041, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT FILING DATE: 2000-01-29

; PRIOR APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-29

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 10041

; LENGTH: 522

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-10041

Query Match 89.3%; Score 25; DB 4; Length 522;

Best Local Similarity 66.7%; Pred. No. 5e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6

Db 387 GYMIEE 392

RESULT 57

US-09-489-039A-7727

; Sequence 7727, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT FILING DATE: 2000-01-29

; PRIOR APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-29

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 10041

; LENGTH: 522

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-7727

Query Match 89.3%; Score 25; DB 4; Length 522;

Best Local Similarity 66.7%; Pred. No. 5e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6

Db 387 GYMIEE 392

RESULT 58

US-09-252-991A-18702

; Sequence 18702, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 18702

; LENGTH: 889

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-18702

Query Match 89.3%; Score 25; DB 4; Length 889;

Best Local Similarity 66.7%; Pred. No. 8.9e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6

Db 714 GYEIEE 719

RESULT 59

US-08-943-768-2

; Sequence 2, Application US/08943768C

; Patent No. 6238881

; GENERAL INFORMATION:

; APPLICANT: Hart, Matthew J.

; TITLE OF INVENTION: No. 6238881el Nucleic Acids and Polypeptides Related to a

; TITLE OF INVENTION: Guanine Exchange Factor of RHO GTPase

; FILE REFERENCE: 1023-US

; CURRENT FILING DATE: 1997-10-06

; EARLIER APPLICATION NUMBER: 60/029,979

; EARLIER FILING DATE: 1996-11-06

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 2

; LENGTH: 912

; TYPE: PRT

; ORGANISM: Human p115 GEF-Rho

US-08-943-768-2

Query Match 89.3%; Score 25; DB 3; Length 912;

```

Best Local Similarity 66.7%; Pred. No. 9.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 486 GYLIEE 491

RESULT 60
US-09-865-960-2
; Sequence 2, Application US/09865960
; Patent No. 6569655
; GENERAL INFORMATION:
; APPLICANT: Hart, Matthew J.
; TITLE OF INVENTION: No. 6569655el Nucleic Acids and Polypeptides Related to a Guanine
; FILE REFERENCE: ONYX1023-DIV1
; CURRENT APPLICATION NUMBER: US/09/865,960
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 08/943,768
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: US 60/029,979
; PRIOR FILING DATE: 1996-11-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Human p115 GEF-Rho gene
; FEATURE:
US-09-865-960-2

Query Match 89.3%; Score 25; DB 4; Length 912;
Best Local Similarity 66.7%; Pred. No. 9.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 486 GYLIEE 491

RESULT 61
US-08-407-804-16
; Sequence 16, Application US/08407804
; Patent No. 5817759
; GENERAL INFORMATION:
; APPLICANT: Margolskee, Robert F.
; TITLE OF INVENTION: Gustducin Materials and Methods
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,804
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,804
; FILING DATE:
; APPLICATION NUMBER: US 07/868/353
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6008000and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-407-804-16

Query Match 85.7%; Score 24; DB 3; Length 13;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-407-804-16

Query Match 85.7%; Score 24; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 4 GYSLEE 9

RESULT 62
US-09-124-807-16
; Sequence 16, Application US/09124807
; Patent No. 6008000
; GENERAL INFORMATION:
; APPLICANT: Margolskee, Robert F.
; TITLE OF INVENTION: Gustducin Materials and Methods
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,807
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,804
; FILING DATE:
; APPLICATION NUMBER: US 07/868/353
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6008000and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-124-807-16

Query Match 85.7%; Score 24; DB 3; Length 13;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```



```
Qy 1 GYXVEE 6
  ||:||
Db 4 GYSLEE 9

RESULT 63
US-09-513-999C-5989
; Sequence 5989, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5989
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5989

Query Match 85.7%; Score 24; DB 4; Length 60;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
  ||:||
Db 13 GYSVDE 18

RESULT 64
US-09-621-976-6862
; Sequence 6862, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6862
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6862

Query Match 85.7%; Score 24; DB 4; Length 64;
Best Local Similarity 66.7%; Pred. No. 91;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
  ||:||
Db 13 GYSVDE 18

RESULT 65
US-09-603-448-13
; Sequence 13, Application US/09603448
; Patent No. 6667153
; GENERAL INFORMATION:
; APPLICANT: Thomas, Susan Margaret
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR DETECTING MUTAGENS
```

```
; FILE REFERENCE: 10552.26US01
; CURRENT APPLICATION NUMBER: US/09/603,448
; CURRENT FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Enterobacteriaceae
US-09-603-448-13

Query Match 85.7%; Score 24; DB 4; Length 74;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
  ||:||
Db 68 GYTVDE 73

RESULT 66
US-09-134-000C-6116
; Sequence 6116, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6116
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6116

Query Match 85.7%; Score 24; DB 4; Length 96;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
  ||:||
Db 70 GYSLEE 75

RESULT 67
US-09-134-001C-3863
; Sequence 3863, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3863
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3863

Query Match 85.7%; Score 24; DB 3; Length 101;
```

Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
||| :||  
Db 43 GYTLEE 48

## RESULT 68

US-09-732-210-85  
; Sequence 85, Application US/09732210  
; Patent No. 6573361  
; GENERAL INFORMATION:  
; APPLICANT: Bunkers, Greg J.  
; APPLICANT: Liang, Jihong  
; APPLICANT: Mittanck, Cindy A.  
; APPLICANT: Seale, Jeffrey W.  
; APPLICANT: Wu, Yonnie S.  
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
; FILE REFERENCE: 38-21(15036)B  
; CURRENT APPLICATION NUMBER: US/09/732,210  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,513  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 85  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Archaeoglobus fulgidus  
US-09-732-210-85

Query Match 85.7%; Score 24; DB 4; Length 133;  
Best Local Similarity 66.7%; Pred. No. 2e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
||| :||  
Db 17 GYTLEE 22

## RESULT 69

US-09-732-210-1505  
; Sequence 1505, Application US/09732210  
; Patent No. 6573361  
; GENERAL INFORMATION:  
; APPLICANT: Bunkers, Greg J.  
; APPLICANT: Liang, Jihong  
; APPLICANT: Mittanck, Cindy A.  
; APPLICANT: Seale, Jeffrey W.  
; APPLICANT: Wu, Yonnie S.  
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
; FILE REFERENCE: 38-21(15036)B  
; CURRENT APPLICATION NUMBER: US/09/732,210  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,513  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 1505  
; LENGTH: 136  
; TYPE: PRT  
; ORGANISM: Methanobacterium thermoautotrophicum  
US-09-732-210-1505

Query Match 85.7%; Score 24; DB 4; Length 136;  
Best Local Similarity 66.7%; Pred. No. 2e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
||| :||

Db 10 GYTLEE 15

## RESULT 70

US-09-710-279-1714  
; Sequence 1714, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P03480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1714  
; LENGTH: 143  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-09-710-279-1714

Query Match 85.7%; Score 24; DB 4; Length 143;  
Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
||| :||  
Db 42 GYTLEE 47

## RESULT 71

US-09-732-210-1504  
; Sequence 1504, Application US/09732210  
; Patent No. 6573361  
; GENERAL INFORMATION:  
; APPLICANT: Bunkers, Greg J.  
; APPLICANT: Liang, Jihong  
; APPLICANT: Mittanck, Cindy A.  
; APPLICANT: Seale, Jeffrey W.  
; APPLICANT: Wu, Yonnie S.  
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
; FILE REFERENCE: 38-21(15036)B  
; CURRENT APPLICATION NUMBER: US/09/732,210  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,513  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 1504  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Methanococcus jannaschii  
US-09-732-210-1504

Query Match 85.7%; Score 24; DB 4; Length 152;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
||| :||  
Db 26 GYTLEE 31

## RESULT 72

US-09-252-991A-22894  
; Sequence 22894, Application US/09252991A  
; Patent No. 6551795

```
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22894
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22894

Query Match      85.7%; Score 24; DB 4; Length 194;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      ||:||
Db      96 GYALEE 101

RESULT 73
US-09-134-000C-4202
; Sequence 4202, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4202
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4202

Query Match      85.7%; Score 24; DB 4; Length 200;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      ||:||
Db      165 GYTL EE 170

RESULT 74
US-09-710-279-776
; Sequence 776, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 776
; LENGTH: 264
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-776

Query Match      85.7%; Score 24; DB 4; Length 264;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      ||:||
Db      49 GYALEE 54

RESULT 75
US-09-134-001C-4729
; Sequence 4729, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4729
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4729

Query Match      85.7%; Score 24; DB 3; Length 268;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
      ||:||
Db      53 GYALEE 58

Search completed: November 1, 2004, 21:35:51
Job time : 42 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:35:11 ; Search time 87 Seconds  
(without alignments)  
22.360 Million cell updates/sec

Title: US-10-030-194A-5  
Perfect score: 28  
Sequence: 1 GYXVEE 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	96.4	151	US-10-425-114-55451	Sequence 56304, A
2	27	96.4	216	US-10-767-701-35832	Sequence 35832, A
3	27	96.4	279	US-10-369-493-9674	Sequence 9674, A
4	27	96.4	282	US-10-136-253-7	Sequence 7, Appli
5	27	96.4	313	US-10-369-493-122	Sequence 122, App
6	27	96.4	344	US-09-815-242-5059	Sequence 5059, App
7	27	96.4	344	US-10-282-122A-43511	Sequence 43511, A
8	27	96.4	365	US-10-369-493-16467	Sequence 16467, A
9	27	96.4	390	US-10-369-493-6541	Sequence 6541, Ap
10	27	96.4	399	US-10-425-114-37084	Sequence 37084, A
11	27	96.4	422	US-10-437-963-117783	Sequence 117783, A
12	27	96.4	456	US-10-369-493-12764	Sequence 12764, A
13	27	96.4	474	US-10-354-437-32	Sequence 32, Appli

14	27	96.4	508	US-10-425-114-55451	Sequence 55451, A
15	27	96.4	524	US-10-424-599-220386	Sequence 220386, A
16	27	96.4	545	US-09-815-242-5591	Sequence 5591, Ap
17	27	96.4	553	US-09-815-242-12198	Sequence 12198, A
18	27	96.4	553	US-10-282-122A-44061	Sequence 44061, A
19	27	96.4	575	US-10-437-963-173223	Sequence 173223, A
20	27	96.4	653	US-10-389-566-1364	Sequence 1364, Ap
21	27	96.4	717	US-10-369-493-9757	Sequence 9757, Ap
22	27	96.4	742	US-10-369-493-18955	Sequence 18955, A
23	27	96.4	762	US-09-391-340-10	Sequence 10, Appl
24	27	96.4	762	US-09-948-369-10	Sequence 10, Appl
25	27	96.4	1075	US-10-416-330-24	Sequence 24, Appl
26	27	96.4	1405	US-10-282-122A-46029	Sequence 46029, A
27	26	92.9	59	US-10-424-599-237501	Sequence 237501, A
28	26	92.9	67	US-10-424-599-212021	Sequence 212021, A
29	26	92.9	84	US-09-864-761-41934	Sequence 41934, A
30	26	92.9	87	US-10-767-701-40109	Sequence 40109, A
31	26	92.9	90	US-09-864-761-42006	Sequence 42006, A
32	26	92.9	95	US-10-335-977-8299	Sequence 8299, Ap
33	26	92.9	102	US-09-333-261-8	Sequence 8, Appli
34	26	92.9	102	US-10-256-702-8	Sequence 8, Appli
35	26	92.9	107	US-10-424-599-268557	Sequence 268557, A
36	26	92.9	112	US-09-864-408A-3458	Sequence 3458, Ap
37	26	92.9	115	US-10-027-450-13	Sequence 13, Appl
38	26	92.9	118	US-10-410-681-20	Sequence 20, Appl
39	26	92.9	118	US-10-424-599-155357	Sequence 155357, A
40	26	92.9	119	US-10-335-977-8300	Sequence 8300, Ap
41	26	92.9	123	US-10-424-599-272017	Sequence 272017, A
42	26	92.9	132	US-10-424-599-159938	Sequence 159938, A
43	26	92.9	139	US-09-864-761-34885	Sequence 34885, A
44	26	92.9	142	US-10-029-386-30849	Sequence 30849, A
45	26	92.9	143	US-10-424-599-169951	Sequence 169951, A
46	26	92.9	145	US-10-424-599-216365	Sequence 216365, A
47	26	92.9	146	US-10-424-599-198169	Sequence 198169, A
48	26	92.9	146	US-10-767-701-40676	Sequence 40676, A
49	26	92.9	154	US-09-925-300-1599	Sequence 1599, Ap
50	26	92.9	160	US-10-276-774-1538	Sequence 1538, Ap
51	26	92.9	163	US-10-424-599-148963	Sequence 148963, A
52	26	92.9	170	US-10-424-599-218619	Sequence 218619, A
53	26	92.9	173	US-10-029-386-33430	Sequence 33430, A
54	26	92.9	181	US-10-437-963-195791	Sequence 195791, A
55	26	92.9	196	US-10-424-599-150556	Sequence 150556, A
56	26	92.9	217	US-10-424-599-203252	Sequence 203252, A
57	26	92.9	225	US-10-424-599-150557	Sequence 150557, A
58	26	92.9	253	US-10-425-114-40929	Sequence 40929, A
59	26	92.9	254	US-10-424-599-158869	Sequence 158869, A
60	26	92.9	255	US-10-424-599-150562	Sequence 150562, A
61	26	92.9	260	US-10-410-681-22	Sequence 22, Appl
62	26	92.9	262	US-10-767-701-43739	Sequence 43739, A
63	26	92.9	263	US-10-425-114-55122	Sequence 55122, A
64	26	92.9	267	US-10-424-599-150560	Sequence 150560, A
65	26	92.9	267	US-10-425-114-71978	Sequence 71978, A
66	26	92.9	267	US-10-425-114-71979	Sequence 71979, A
67	26	92.9	271	US-10-425-114-52770	Sequence 52770, A
68	26	92.9	278	US-10-424-599-271699	Sequence 271699, A
69	26	92.9	286	US-10-029-386-33440	Sequence 33440, A
70	26	92.9	289	US-10-425-114-40821	Sequence 40821, A
71	26	92.9	303	US-10-156-761-11627	Sequence 11627, A
72	26	92.9	307	US-10-027-450-15	Sequence 15, Appl
73	26	92.9	319	US-10-425-114-37774	Sequence 37774, A
74	26	92.9	326	US-09-988-626-231	Sequence 231, App
75	26	92.9	326	US-09-988-687-231	Sequence 231, App
76	26	92.9	326	US-09-988-686-231	Sequence 231, App
77	26	92.9	327	US-10-282-122A-51411	Sequence 51411, A
78	26	92.9	327	US-10-425-114-37141	Sequence 37141, A
79	26	92.9	345	US-10-369-493-16642	Sequence 16642, A
80	26	92.9	348	US-10-369-493-8868	Sequence 8868, Ap
81	26	92.9	358	US-10-282-122A-70250	Sequence 70250, A
82	26	92.9	358	US-10-282-122A-70471	Sequence 70471, A
83	26	92.9	358	US-10-282-122A-71576	Sequence 71576, A
84	26	92.9	362	US-10-369-493-3080	Sequence 3080, Ap
85	26	92.9	363	US-10-425-114-67820	Sequence 67820, A
86	26	92.9	367	US-10-020-513-2	Sequence 2, Appli

87	26	92.9	370	14	US-10-369-493-11308	Sequence 11308, A	160	26	92.9	945	14	US-10-123-261-146	Sequence 146, App
88	26	92.9	379	9	US-09-738-626-5914	Sequence 5914, Ap	161	26	92.9	945	14	US-10-140-921-146	Sequence 146, App
89	26	92.9	381	15	US-10-425-114-38181	Sequence 38181, A	162	26	92.9	945	14	US-10-140-928-146	Sequence 146, App
90	26	92.9	388	15	US-10-424-598-156035	Sequence 156035,	163	26	92.9	945	14	US-10-121-045-146	Sequence 146, App
91	26	92.9	386	10	US-09-374-046A-42	Sequence 42, Appl	164	26	92.9	945	14	US-10-123-292-146	Sequence 146, App
92	26	92.9	396	15	US-10-616-263-42	Sequence 42, Appl	165	26	92.9	945	14	US-10-123-903-146	Sequence 146, App
93	26	92.9	410	14	US-10-369-493-19854	Sequence 19854, A	166	26	92.9	945	14	US-10-124-819-146	Sequence 146, App
94	26	92.9	413	14	US-10-225-066A-1018	Sequence 1018, Ap	167	26	92.9	945	14	US-10-124-822-146	Sequence 146, App
95	26	92.9	413	15	US-10-374-780A-2712	Sequence 2712, Ap	168	26	92.9	945	14	US-10-140-925-146	Sequence 146, App
96	26	92.9	425	15	US-10-425-114-44761	Sequence 44761, A	169	26	92.9	945	14	US-10-160-498-146	Sequence 146, App
97	26	92.9	458	15	US-10-296-115-804	Sequence 804, App	170	26	92.9	945	14	US-10-124-824-146	Sequence 146, App
98	26	92.9	484	16	US-10-437-963-171845	Sequence 171845,	171	26	92.9	945	14	US-10-127-825A-146	Sequence 146, App
99	26	92.9	523	15	US-10-424-599-151286	Sequence 151286,	172	26	92.9	945	14	US-10-127-829A-146	Sequence 146, App
100	26	92.9	532	9	US-09-911-513-2	Sequence 2, Appli	173	26	92.9	945	14	US-10-127-835A-146	Sequence 146, App
101	26	92.9	532	10	US-09-911-514-2	Sequence 2, Appli	174	26	92.9	945	14	US-10-127-838A-146	Sequence 146, App
102	26	92.9	532	14	US-10-278-536-46	Sequence 46, Appl	175	26	92.9	945	14	US-10-127-901A-146	Sequence 146, App
103	26	92.9	532	14	US-10-225-086A-784	Sequence 784, App	176	26	92.9	945	14	US-10-127-901A-146	Sequence 146, App
104	26	92.9	532	15	US-10-374-780A-2290	Sequence 2290, Ap	177	26	92.9	945	14	US-10-128-693A-146	Sequence 146, App
105	26	92.9	532	15	US-10-412-699B-218	Sequence 218, App	177	26	92.9	945	14	US-10-131-813A-146	Sequence 146, App
106	26	92.9	534	15	US-10-282-122A-70794	Sequence 70794, A	178	26	92.9	945	14	US-10-131-818A-146	Sequence 146, App
107	26	92.9	544	15	US-10-425-114-42064	Sequence 42064, A	179	26	92.9	945	14	US-10-131-823A-146	Sequence 146, App
108	26	92.9	546	15	US-10-425-114-72866	Sequence 72866, A	180	26	92.9	945	14	US-10-131-824A-146	Sequence 146, App
109	26	92.9	554	15	US-10-108-260A-2682	Sequence 2682, Ap	181	26	92.9	945	14	US-10-131-830A-146	Sequence 146, App
110	26	92.9	578	16	US-10-437-963-195324	Sequence 195324,	182	26	92.9	945	14	US-10-131-837A-146	Sequence 146, App
111	26	92.9	587	15	US-10-412-699B-216	Sequence 216, App	183	26	92.9	945	14	US-10-137-872A-146	Sequence 146, App
112	26	92.9	595	15	US-10-424-599-268568	Sequence 268568,	184	26	92.9	945	14	US-10-147-500-146	Sequence 146, App
113	26	92.9	595	16	US-10-437-963-176083	Sequence 176083,	185	26	92.9	945	14	US-10-147-502-146	Sequence 146, App
114	26	92.9	606	14	US-10-369-493-6673	Sequence 6673, Ap	186	26	92.9	945	14	US-10-147-515-146	Sequence 146, App
115	26	92.9	614	14	US-10-369-493-14249	Sequence 14249, A	187	26	92.9	945	14	US-10-147-517-146	Sequence 146, App
116	26	92.9	625	14	US-10-310-154-562	Sequence 562, App	188	26	92.9	945	14	US-10-147-526-146	Sequence 146, App
117	26	92.9	630	14	US-10-310-154-569	Sequence 569, App	189	26	92.9	945	14	US-10-147-527-146	Sequence 146, App
118	26	92.9	698	15	US-10-351-334-200	Sequence 200, App	190	26	92.9	945	14	US-10-121-041-146	Sequence 146, App
119	26	92.9	712	16	US-10-781-014-596	Sequence 596, App	191	26	92.9	945	14	US-10-121-043-146	Sequence 146, App
120	26	92.9	739	9	US-09-738-626-6988	Sequence 6988, Ap	192	26	92.9	945	14	US-10-121-047-146	Sequence 146, App
121	26	92.9	739	16	US-10-781-014-594	Sequence 594, App	193	26	92.9	945	14	US-10-123-215-146	Sequence 146, App
122	26	92.9	752	14	US-10-104-047-2003	Sequence 3003, Ap	194	26	92.9	945	14	US-10-123-902-146	Sequence 146, App
123	26	92.9	809	14	US-10-369-493-15146	Sequence 15146, A	195	26	92.9	945	14	US-10-123-908-146	Sequence 146, App
124	26	92.9	810	14	US-10-369-493-11716	Sequence 11716, A	196	26	92.9	945	14	US-10-123-909-146	Sequence 146, App
125	26	92.9	820	16	US-10-437-963-140417	Sequence 140417,	197	26	92.9	945	14	US-10-123-910-146	Sequence 146, App
126	26	92.9	824	16	US-10-437-963-171380	Sequence 171380,	198	26	92.9	945	14	US-10-124-813-146	Sequence 146, App
127	26	92.9	865	15	US-10-282-122A-69151	Sequence 69151, A	199	26	92.9	945	14	US-10-124-817-146	Sequence 146, App
128	26	92.9	873	13	US-10-167-361-2	Sequence 2, Appli	200	26	92.9	945	14	US-10-125-922-146	Sequence 146, App
129	26	92.9	873	16	US-10-617-351-9	Sequence 9, Appli	201	26	92.9	945	14	US-10-125-924-146	Sequence 146, App
130	26	92.9	873	17	US-10-479-875-7	Sequence 7, Appli	202	26	92.9	945	14	US-10-140-860-146	Sequence 146, App
131	26	92.9	933	15	US-10-087-684-2	Sequence 2, Appli	203	26	92.9	945	14	US-10-142-417-146	Sequence 146, App
132	26	92.9	933	15	US-10-087-684-2	Sequence 2, Appli	204	26	92.9	945	14	US-10-157-782-146	Sequence 146, App
133	26	92.9	933	15	US-10-218-779-2	Sequence 2, Appli	205	26	92.9	945	14	US-10-152-385-146	Sequence 146, App
134	26	92.9	933	15	US-10-218-779-2	Sequence 4, Appli	206	26	92.9	945	14	US-10-125-924-146	Sequence 146, App
135	26	92.9	943	10	US-09-933-261-7	Sequence 7, Appli	207	26	92.9	945	14	US-10-127-838B-146	Sequence 146, App
136	26	92.9	943	14	US-10-256-702-7	Sequence 7, Appli	208	26	92.9	945	14	US-10-127-842A-146	Sequence 146, App
137	26	92.9	945	11	US-09-972-211-123	Sequence 123, App	209	26	92.9	945	14	US-10-127-843A-146	Sequence 146, App
138	26	92.9	945	11	US-09-972-211-124	Sequence 124, App	210	26	92.9	945	14	US-10-127-837A-146	Sequence 146, App
139	26	92.9	945	14	US-10-028-072-146	Sequence 146, App	211	26	92.9	945	14	US-10-127-838B-146	Sequence 146, App
140	26	92.9	945	14	US-10-140-808-146	Sequence 146, App	212	26	92.9	945	14	US-10-127-842A-146	Sequence 146, App
141	26	92.9	945	14	US-10-121-049-146	Sequence 146, App	213	26	92.9	945	14	US-10-127-845A-146	Sequence 146, App
142	26	92.9	945	14	US-10-123-904-146	Sequence 146, App	214	26	92.9	945	14	US-10-127-845A-146	Sequence 146, App
143	26	92.9	945	14	US-10-140-470-146	Sequence 146, App	215	26	92.9	945	14	US-10-127-846A-146	Sequence 146, App
144	26	92.9	945	14	US-10-175-746-146	Sequence 146, App	216	26	92.9	945	14	US-10-127-848A-146	Sequence 146, App
145	26	92.9	945	14	US-10-176-918-146	Sequence 146, App	217	26	92.9	945	14	US-10-127-849A-146	Sequence 146, App
146	26	92.9	945	14	US-10-176-921-146	Sequence 146, App	218	26	92.9	945	14	US-10-127-850A-146	Sequence 146, App
147	26	92.9	945	14	US-10-137-865-146	Sequence 146, App	219	26	92.9	945	14	US-10-127-851A-146	Sequence 146, App
148	26	92.9	945	14	US-10-140-474-146	Sequence 146, App	220	26	92.9	945	14	US-10-128-684A-146	Sequence 146, App
149	26	92.9	945	14	US-10-142-423-146	Sequence 146, App	221	26	92.9	945	14	US-10-128-686A-146	Sequence 146, App
150	26	92.9	945	14	US-10-142-431-146	Sequence 146, App	222	26	92.9	945	14	US-10-128-690A-146	Sequence 146, App
151	26	92.9	945	14	US-10-143-114-146	Sequence 146, App	223	26	92.9	945	14	US-10-128-691A-146	Sequence 146, App
152	26	92.9	945	14	US-10-140-002-146	Sequence 146, App	224	26	92.9	945	14	US-10-131-819A-146	Sequence 146, App
153	26	92.9	945	14	US-10-142-419-146	Sequence 146, App	225	26	92.9	945	14	US-10-131-829A-146	Sequence 146, App
154	26	92.9	945	14	US-10-123-262-146	Sequence 146, App	226	26	92.9	945	14	US-10-131-829A-146	Sequence 146, App
155	26	92.9	945	14	US-10-142-423-146	Sequence 146, App	227	26	92.9	945	14	US-10-136-729-146	Sequence 146, App
156	26	92.9	945	14	US-10-121-050-146	Sequence 146, App	228	26	92.9	945	14	US-10-146-791-146	Sequence 146, App
157	26	92.9	945	14	US-10-141-755-146	Sequence 146, App	229	26	92.9	945	14	US-10-147-484-146	Sequence 146, App
158	26	92.9	945	14	US-10-143-032-146	Sequence 146, App	230	26	92.9	945	14	US-10-147-508-146	Sequence 146, App
159	26	92.9	945	14	US-10-123-108-146	Sequence 146, App	231	26	92.9	945	14	US-10-147-512-146	Sequence 146, App
	26	92.9	945	14	US-10-123-236-146	Sequence 146, App	232	26	92.9	945	14	US-10-175-735-146	Sequence 146, App

233	26	92.9	945	14	US-10-121-040-146	Sequence 146, App	306
234	26	92.9	945	14	US-10-121-056-146	Sequence 146, App	307
235	26	92.9	945	14	US-10-121-061-146	Sequence 146, App	308
236	26	92.9	945	14	US-10-123-235-146	Sequence 146, App	309
237	26	92.9	945	14	US-10-124-818-146	Sequence 146, App	310
238	26	92.9	945	14	US-10-137-868-146	Sequence 146, App	311
239	26	92.9	945	14	US-10-147-492-146	Sequence 146, App	312
240	26	92.9	945	14	US-10-158-782-146	Sequence 146, App	313
241	26	92.9	945	14	US-10-123-905-146	Sequence 146, App	314
242	26	92.9	945	14	US-10-123-907-146	Sequence 146, App	315
243	26	92.9	945	14	US-10-124-815-146	Sequence 146, App	316
244	26	92.9	945	14	US-10-125-821A-146	Sequence 146, App	317
245	26	92.9	945	14	US-10-125-928A-146	Sequence 146, App	318
246	26	92.9	945	14	US-10-127-821A-146	Sequence 146, App	319
247	26	92.9	945	14	US-10-127-822A-146	Sequence 146, App	320
248	26	92.9	945	14	US-10-127-833A-146	Sequence 146, App	321
249	26	92.9	945	14	US-10-127-826A-146	Sequence 146, App	322
250	26	92.9	945	14	US-10-127-827A-146	Sequence 146, App	323
251	26	92.9	945	14	US-10-127-828A-146	Sequence 146, App	324
252	26	92.9	945	14	US-10-127-830A-146	Sequence 146, App	325
253	26	92.9	945	14	US-10-127-832A-146	Sequence 146, App	326
254	26	92.9	945	14	US-10-127-833A-146	Sequence 146, App	327
255	26	92.9	945	14	US-10-127-834A-146	Sequence 146, App	328
256	26	92.9	945	14	US-10-127-836A-146	Sequence 146, App	329
257	26	92.9	945	14	US-10-127-841A-146	Sequence 146, App	330
258	26	92.9	945	14	US-10-127-844A-146	Sequence 146, App	331
259	26	92.9	945	14	US-10-128-687A-146	Sequence 146, App	332
260	26	92.9	945	14	US-10-128-688A-146	Sequence 146, App	333
261	26	92.9	945	14	US-10-128-689A-146	Sequence 146, App	334
262	26	92.9	945	14	US-10-128-694A-146	Sequence 146, App	335
263	26	92.9	945	14	US-10-131-825A-146	Sequence 146, App	336
264	26	92.9	945	14	US-10-230-417-146	Sequence 146, App	337
265	26	92.9	945	14	US-10-131-815A-146	Sequence 146, App	338
266	26	92.9	945	14	US-10-131-817A-146	Sequence 146, App	339
267	26	92.9	945	14	US-10-131-821A-146	Sequence 146, App	340
268	26	92.9	945	14	US-10-131-822A-146	Sequence 146, App	341
269	26	92.9	945	14	US-10-131-828A-146	Sequence 146, App	342
270	26	92.9	945	14	US-10-131-835A-146	Sequence 146, App	343
271	26	92.9	945	14	US-10-137-864A-146	Sequence 146, App	344
272	26	92.9	945	14	US-10-137-869A-146	Sequence 146, App	345
273	26	92.9	945	14	US-10-147-523-146	Sequence 146, App	346
274	26	92.9	945	14	US-10-158-785-146	Sequence 146, App	347
275	26	92.9	945	14	US-10-121-051-146	Sequence 146, App	348
276	26	92.9	945	14	US-10-121-042-146	Sequence 146, App	349
277	26	92.9	945	14	US-10-123-912-146	Sequence 146, App	350
278	26	92.9	945	14	US-10-132-007-146	Sequence 146, App	351
279	26	92.9	945	14	US-10-134-359-146	Sequence 146, App	352
280	26	92.9	945	14	US-10-127-847A-146	Sequence 146, App	353
281	26	92.9	945	14	US-10-175-590-146	Sequence 146, App	354
282	26	92.9	945	14	US-10-153-934-146	Sequence 146, App	355
283	26	92.9	945	14	US-10-146-726-146	Sequence 146, App	356
284	26	92.9	945	14	US-10-146-727-146	Sequence 146, App	357
285	26	92.9	945	14	US-10-146-788-146	Sequence 146, App	358
286	26	92.9	945	14	US-10-152-380-146	Sequence 146, App	359
287	26	92.9	945	14	US-10-153-934-146	Sequence 146, App	360
288	26	92.9	945	14	US-10-140-807-146	Sequence 146, App	361
289	26	92.9	945	14	US-10-140-924-146	Sequence 146, App	362
290	26	92.9	945	14	US-10-140-926-146	Sequence 146, App	363
291	26	92.9	945	14	US-10-141-698-146	Sequence 146, App	364
292	26	92.9	945	14	US-10-141-702-146	Sequence 146, App	365
293	26	92.9	945	14	US-10-141-704-146	Sequence 146, App	366
294	26	92.9	945	14	US-10-142-432-146	Sequence 146, App	367
295	26	92.9	945	14	US-10-142-421-146	Sequence 146, App	368
296	26	92.9	945	14	US-10-142-767-146	Sequence 146, App	369
297	26	92.9	945	14	US-10-143-033-146	Sequence 146, App	370
298	26	92.9	945	14	US-10-144-994-146	Sequence 146, App	371
299	26	92.9	945	14	US-10-145-628-146	Sequence 146, App	372
300	26	92.9	945	14	US-10-145-746-146	Sequence 146, App	373
301	26	92.9	945	14	US-10-145-748-146	Sequence 146, App	374
302	26	92.9	945	14	US-10-145-823-146	Sequence 146, App	375
303	26	92.9	945	14	US-10-145-826-146	Sequence 146, App	376
304	26	92.9	945	14	US-10-145-870-146	Sequence 146, App	377
305	26	92.9	945	14	US-10-145-876-146	Sequence 146, App	378
306	26	92.9	945	14	US-10-145-959-146	Sequence 146, App	379
307	26	92.9	945	14	US-10-146-724-146	Sequence 146, App	380
308	26	92.9	945	14	US-10-146-725-146	Sequence 146, App	381
309	26	92.9	945	14	US-10-146-795-146	Sequence 146, App	382
310	26	92.9	945	14	US-10-147-495-146	Sequence 146, App	383
311	26	92.9	945	14	US-10-147-501-146	Sequence 146, App	384
312	26	92.9	945	14	US-10-147-504-146	Sequence 146, App	385
313	26	92.9	945	14	US-10-147-506-146	Sequence 146, App	386
314	26	92.9	945	14	US-10-147-509-146	Sequence 146, App	387
315	26	92.9	945	14	US-10-147-510-146	Sequence 146, App	388
316	26	92.9	945	14	US-10-147-511-146	Sequence 146, App	389
317	26	92.9	945	14	US-10-147-529-146	Sequence 146, App	390
318	26	92.9	945	14	US-10-152-397-146	Sequence 146, App	391
319	26	92.9	945	14	US-10-153-586-146	Sequence 146, App	392
320	26	92.9	945	14	US-10-158-786-146	Sequence 146, App	393
321	26	92.9	945	14	US-10-137-870-146	Sequence 146, App	394
322	26	92.9	945	14	US-10-140-018-146	Sequence 146, App	395
323	26	92.9	945	14	US-10-140-021-146	Sequence 146, App	396
324	26	92.9	945	14	US-10-140-019-146	Sequence 146, App	397
325	26	92.9	945	14	US-10-140-922-146	Sequence 146, App	398
326	26	92.9	945	14	US-10-140-861-146	Sequence 146, App	399
327	26	92.9	945	14	US-10-140-862-146	Sequence 146, App	400
328	26	92.9	945	14	US-10-140-862-146	Sequence 146, App	401
329	26	92.9	945	14	US-10-141-697-146	Sequence 146, App	402
330	26	92.9	945	14	US-10-141-700-146	Sequence 146, App	403
331	26	92.9	945	14	US-10-141-705-146	Sequence 146, App	404
332	26	92.9	945	14	US-10-141-753-146	Sequence 146, App	405
333	26	92.9	945	14	US-10-141-758-146	Sequence 146, App	406
334	26	92.9	945	14	US-10-142-418-146	Sequence 146, App	407
335	26	92.9	945	14	US-10-142-420-146	Sequence 146, App	408
336	26	92.9	945	14	US-10-142-422-146	Sequence 146, App	409
337	26	92.9	945	14	US-10-142-427-146	Sequence 146, App	410
338	26	92.9	945	14	US-10-142-760-146	Sequence 146, App	411
339	26	92.9	945	14	US-10-145-821-146	Sequence 146, App	412
340	26	92.9	945	14	US-10-152-531-146	Sequence 146, App	413
341	26	92.9	945	14	US-10-127-840A-146	Sequence 146, App	414
342	26	92.9	945	14	US-10-142-424-146	Sequence 146, App	415
343	26	92.9	945	14	US-10-142-761-146	Sequence 146, App	416
344	26	92.9	945	14	US-10-142-763-146	Sequence 146, App	417
345	26	92.9	945	14	US-10-142-765-146	Sequence 146, App	418
346	26	92.9	945	14	US-10-142-887-146	Sequence 146, App	419
347	26	92.9	945	14	US-10-142-888-146	Sequence 146, App	420
348	26	92.9	945	14	US-10-143-034-146	Sequence 146, App	421
349	26	92.9	945	14	US-10-143-116-146	Sequence 146, App	422
350	26	92.9	945	14	US-10-145-015-146	Sequence 146, App	423
351	26	92.9	945	14	US-10-145-090-146	Sequence 146, App	424
352	26	92.9	945	14	US-10-145-091-146	Sequence 146, App	425
353	26	92.9	945	14	US-10-145-629-146	Sequence 146, App	426
354	26	92.9	945	14	US-10-145-630-146	Sequence 146, App	427
355	26	92.9	945	14	US-10-144-992-146	Sequence 146, App	428
356	26	92.9	945	14	US-10-144-992-146	Sequence 146, App	429
357	26	92.9	945	14	US-10-145-015-146	Sequence 146, App	430
358	26	92.9	945	14	US-10-145-090-146	Sequence 146, App	431
359	26	92.9	945	14	US-10-145-091-146	Sequence 146, App	432
360	26	92.9	945	14	US-10-145-629-146	Sequence 146, App	433
361	26	92.9	945	14	US-10-145-630-146	Sequence 146, App	434
362	26	92.9	945	14	US-10-145-747-146	Sequence 146, App	435
363	26	92.9	945	14	US-10-145-752-146	Sequence 146, App	436
364	26	92.9	945	14	US-10-145-754-146	Sequence 146, App	437
365	26	92.9	945	14	US-10-145-755-146	Sequence 146, App	438
366	26	92.9	945	14	US-10-145-818-146	Sequence 146, App	439
367	26	92.9	945	14	US-10-145-820-146	Sequence 146, App	440
368	26	92.9	945	14	US-10-145-872-146	Sequence 146, App	441
369	26	92.9	945	14	US-10-145-873-146	Sequence 146, App	442
370	26	92.9	945	14	US-10-147-481-146	Sequence 146, App	443
371	26	92.9	945	14	US-10-147-482-146	Sequence 146, App	444
372	26	92.9	945	14	US-10-147-503-146	Sequence 146, App	445
373	26	92.9	945	14	US-10-147-522-146	Sequence 146, App	446
374	26	92.9	945	14	US-10-152-401-146	Sequence 146, App	447
375	26	92.9	945	14	US-10-157-783-146	Sequence 146, App	448
376	26	92.9	945	14	US-10-158-792-146	Sequence 146, App	449
377	26	92.9	945	14	US-10-158-862-146	Sequence 146, App	450
378	26	92.9	945	14	US-10-143-035-146	Sequence 146, App	451
379	26	92.9	945	14	US-10-143-035-146	Sequence 146, App	452

379	26	92.9	945	14	US-10-145-751-146	Sequence 146, App	452	26	92.9	945	14	US-10-124-820-146	Sequence 146, App
380	26	92.9	945	14	US-10-145-822-146	Sequence 146, App	453	26	92.9	945	14	US-10-125-704-146	Sequence 146, App
381	26	92.9	945	14	US-10-145-824-146	Sequence 146, App	454	26	92.9	945	14	US-10-125-927-146	Sequence 146, App
382	26	92.9	945	14	US-10-145-827-146	Sequence 146, App	455	26	92.9	945	14	US-10-142-889-146	Sequence 146, App
383	26	92.9	945	14	US-10-145-869-146	Sequence 146, App	456	26	92.9	945	14	US-10-145-874-146	Sequence 146, App
384	26	92.9	945	14	US-10-145-875-146	Sequence 146, App	457	26	92.9	945	14	US-10-147-497-146	Sequence 146, App
385	26	92.9	945	14	US-10-145-877-146	Sequence 146, App	458	26	92.9	945	14	US-10-152-371-146	Sequence 146, App
386	26	92.9	945	14	US-10-145-898-146	Sequence 146, App	459	26	92.9	945	14	US-10-152-374-146	Sequence 146, App
387	26	92.9	945	14	US-10-146-780-146	Sequence 146, App	460	26	92.9	945	14	US-10-152-375-146	Sequence 146, App
388	26	92.9	945	14	US-10-146-790-146	Sequence 146, App	461	26	92.9	945	14	US-10-152-377-146	Sequence 146, App
389	26	92.9	945	14	US-10-146-793-146	Sequence 146, App	462	26	92.9	945	14	US-10-152-386-146	Sequence 146, App
390	26	92.9	945	14	US-10-147-480-146	Sequence 146, App	463	26	92.9	945	14	US-10-152-391-146	Sequence 146, App
391	26	92.9	945	14	US-10-147-485-146	Sequence 146, App	464	26	92.9	945	14	US-10-152-399-146	Sequence 146, App
392	26	92.9	945	14	US-10-147-486-146	Sequence 146, App	465	26	92.9	945	14	US-10-156-848-146	Sequence 146, App
393	26	92.9	945	14	US-10-147-487-146	Sequence 146, App	466	26	92.9	945	14	US-10-157-785-146	Sequence 146, App
394	26	92.9	945	14	US-10-147-490-146	Sequence 146, App	467	26	92.9	945	14	US-10-157-794-146	Sequence 146, App
395	26	92.9	945	14	US-10-147-494-146	Sequence 146, App	468	26	92.9	945	14	US-10-157-796-146	Sequence 146, App
396	26	92.9	945	14	US-10-147-498-146	Sequence 146, App	469	26	92.9	945	14	US-10-160-500-146	Sequence 146, App
397	26	92.9	945	14	US-10-147-514-146	Sequence 146, App	470	26	92.9	945	14	US-10-121-046-146	Sequence 146, App
398	26	92.9	945	14	US-10-147-524-146	Sequence 146, App	471	26	92.9	945	14	US-10-123-156-146	Sequence 146, App
399	26	92.9	945	14	US-10-152-379-146	Sequence 146, App	472	26	92.9	945	14	US-10-123-214-146	Sequence 146, App
400	26	92.9	945	14	US-10-152-394-146	Sequence 146, App	473	26	92.9	945	14	US-10-125-805-146	Sequence 146, App
401	26	92.9	945	14	US-10-152-406-146	Sequence 146, App	474	26	92.9	945	14	US-10-124-821-146	Sequence 146, App
402	26	92.9	945	14	US-10-156-847-146	Sequence 146, App	475	26	92.9	945	14	US-10-152-385-146	Sequence 146, App
403	26	92.9	945	14	US-10-157-779-146	Sequence 146, App	476	26	92.9	945	14	US-10-152-393-146	Sequence 146, App
404	26	92.9	945	14	US-10-157-788-146	Sequence 146, App	477	26	92.9	945	14	US-10-152-396-146	Sequence 146, App
405	26	92.9	945	14	US-10-160-504-146	Sequence 146, App	478	26	92.9	945	14	US-10-153-552-146	Sequence 146, App
406	26	92.9	945	14	US-10-145-834-146	Sequence 146, App	479	26	92.9	945	14	US-10-153-840-146	Sequence 146, App
407	26	92.9	945	14	US-10-147-520-146	Sequence 146, App	480	26	92.9	945	14	US-10-156-841-146	Sequence 146, App
408	26	92.9	945	14	US-10-157-781-146	Sequence 146, App	481	26	92.9	945	14	US-10-156-842-146	Sequence 146, App
409	26	92.9	945	14	US-10-176-989-146	Sequence 146, App	482	26	92.9	945	14	US-10-156-844-146	Sequence 146, App
410	26	92.9	945	14	US-10-147-491-146	Sequence 146, App	483	26	92.9	945	14	US-10-156-845-146	Sequence 146, App
411	26	92.9	945	14	US-10-152-378-146	Sequence 146, App	484	26	92.9	945	14	US-10-156-846-146	Sequence 146, App
412	26	92.9	945	14	US-10-152-382-146	Sequence 146, App	485	26	92.9	945	14	US-10-121-048-146	Sequence 146, App
413	26	92.9	945	14	US-10-152-383-146	Sequence 146, App	486	26	92.9	945	14	US-10-121-052-146	Sequence 146, App
414	26	92.9	945	14	US-10-152-384-146	Sequence 146, App	487	26	92.9	945	14	US-10-121-053-146	Sequence 146, App
415	26	92.9	945	14	US-10-152-387-146	Sequence 146, App	488	26	92.9	945	14	US-10-121-054-146	Sequence 146, App
416	26	92.9	945	14	US-10-152-389-146	Sequence 146, App	489	26	92.9	945	14	US-10-121-063-146	Sequence 146, App
417	26	92.9	945	14	US-10-152-390-146	Sequence 146, App	490	26	92.9	945	14	US-10-123-212-146	Sequence 146, App
418	26	92.9	945	14	US-10-152-392-146	Sequence 146, App	491	26	92.9	945	14	US-10-123-213-146	Sequence 146, App
419	26	92.9	945	14	US-10-153-756-146	Sequence 146, App	492	26	92.9	945	14	US-10-123-291-146	Sequence 146, App
420	26	92.9	945	14	US-10-157-779-146	Sequence 146, App	493	26	92.9	945	14	US-10-123-322-146	Sequence 146, App
421	26	92.9	945	14	US-10-137-784-146	Sequence 146, App	494	26	92.9	945	14	US-10-123-771-146	Sequence 146, App
422	26	92.9	945	14	US-10-157-797-146	Sequence 146, App	495	26	92.9	945	14	US-10-123-911-146	Sequence 146, App
423	26	92.9	945	14	US-10-158-491-146	Sequence 146, App	496	26	92.9	945	14	US-10-124-823-146	Sequence 146, App
424	26	92.9	945	14	US-10-142-762-146	Sequence 146, App	497	26	92.9	945	14	US-10-125-931-146	Sequence 146, App
425	26	92.9	945	14	US-10-142-764-146	Sequence 146, App	498	26	92.9	945	14	US-10-125-932-146	Sequence 146, App
426	26	92.9	945	14	US-10-142-766-146	Sequence 146, App	499	26	92.9	945	14	US-10-127-852A-146	Sequence 146, App
427	26	92.9	945	14	US-10-145-625-146	Sequence 146, App	500	26	92.9	945	14	US-10-127-900A-146	Sequence 146, App
428	26	92.9	945	14	US-10-145-627-146	Sequence 146, App	501	26	92.9	945	14	US-10-128-685A-146	Sequence 146, App
429	26	92.9	945	14	US-10-145-960-146	Sequence 146, App	502	26	92.9	945	14	US-10-131-820A-146	Sequence 146, App
430	26	92.9	945	14	US-10-145-962-146	Sequence 146, App	503	26	92.9	945	14	US-10-142-886-146	Sequence 146, App
431	26	92.9	945	14	US-10-146-789-146	Sequence 146, App	504	26	92.9	945	14	US-10-146-728-146	Sequence 146, App
432	26	92.9	945	14	US-10-147-483-146	Sequence 146, App	505	26	92.9	945	14	US-10-146-786-146	Sequence 146, App
433	26	92.9	945	14	US-10-147-496-146	Sequence 146, App	506	26	92.9	945	14	US-10-147-499-146	Sequence 146, App
434	26	92.9	945	14	US-10-147-505-146	Sequence 146, App	507	26	92.9	945	14	US-10-157-798-146	Sequence 146, App
435	26	92.9	945	14	US-10-147-516-146	Sequence 146, App	508	26	92.9	945	14	US-10-123-913-146	Sequence 146, App
436	26	92.9	945	14	US-10-152-398-146	Sequence 146, App	509	26	92.9	945	14	US-10-140-473-146	Sequence 146, App
437	26	92.9	945	14	US-10-139-980-146	Sequence 146, App	510	26	92.9	945	14	US-10-140-806-146	Sequence 146, App
438	26	92.9	945	14	US-10-145-750-146	Sequence 146, App	511	26	92.9	945	14	US-10-140-810-146	Sequence 146, App
439	26	92.9	945	14	US-10-152-373-146	Sequence 146, App	512	26	92.9	945	14	US-10-140-863-146	Sequence 146, App
440	26	92.9	945	14	US-10-121-044-146	Sequence 146, App	513	26	92.9	945	14	US-10-141-699-146	Sequence 146, App
441	26	92.9	945	14	US-10-121-055-146	Sequence 146, App	514	26	92.9	945	14	US-10-141-703-146	Sequence 146, App
442	26	92.9	945	14	US-10-121-057-146	Sequence 146, App	515	26	92.9	945	14	US-10-141-706-146	Sequence 146, App
443	26	92.9	945	14	US-10-121-058-146	Sequence 146, App	516	26	92.9	945	14	US-10-141-757-146	Sequence 146, App
444	26	92.9	945	14	US-10-121-059-146	Sequence 146, App	517	26	92.9	945	14	US-10-141-762-146	Sequence 146, App
445	26	92.9	945	14	US-10-121-060-146	Sequence 146, App	518	26	92.9	945	14	US-10-142-428-146	Sequence 146, App
446	26	92.9	945	14	US-10-123-109-146	Sequence 146, App	519	26	92.9	945	14	US-10-142-429-146	Sequence 146, App
447	26	92.9	945	14	US-10-123-154-146	Sequence 146, App	520	26	92.9	945	14	US-10-142-884-146	Sequence 146, App
448	26	92.9	945	14	US-10-123-157-146	Sequence 146, App	521	26	92.9	945	14	US-10-143-027-146	Sequence 146, App
449	26	92.9	945	14	US-10-123-906-146	Sequence 146, App	522	26	92.9	945	14	US-10-143-115-146	Sequence 146, App
450	26	92.9	945	14	US-10-124-814-146	Sequence 146, App	523	26	92.9	945	14	US-10-144-956-146	Sequence 146, App
451	26	92.9	945	14	US-10-124-816-146	Sequence 146, App	524	26	92.9	945	14	US-10-144-958-146	Sequence 146, App



525	26	92.9	945	14	US-10-145-632-146	Sequence 146, App	598	26	92.9	1403	16	US-10-437-963-137174	Sequence 137174,
526	26	92.9	945	14	US-10-145-749-146	Sequence 146, App	599	25	89.3	172	16	US-10-437-963-148744	Sequence 148744,
527	26	92.9	945	14	US-10-145-753-146	Sequence 146, App	600	25	89.3	109	15	US-10-335-977-9348	Sequence 9348, App
528	26	92.9	945	14	US-10-145-871-146	Sequence 146, App	601	25	89.3	113	16	US-10-767-701-57618	Sequence 57618, A
529	26	92.9	945	14	US-10-145-878-146	Sequence 146, App	602	25	89.3	142	16	US-10-437-963-156182	Sequence 156182, A
530	26	92.9	945	14	US-10-146-794-146	Sequence 146, App	603	25	89.3	144	16	US-10-767-701-56458	Sequence 56458, A
531	26	92.9	945	14	US-10-147-489-146	Sequence 146, App	604	25	89.3	170	15	US-10-425-114-46945	Sequence 46945, A
532	26	92.9	945	14	US-10-147-507-146	Sequence 146, App	605	25	89.3	173	15	US-10-425-114-65873	Sequence 65873, A
533	26	92.9	945	14	US-10-147-535-146	Sequence 146, App	606	25	89.3	180	15	US-10-282-122A-59062	Sequence 59062, A
534	26	92.9	945	14	US-10-147-537-146	Sequence 146, App	607	25	89.3	180	15	US-10-335-977-8345	Sequence 8345, App
535	26	92.9	945	14	US-10-152-376-146	Sequence 146, App	608	25	89.3	180	15	US-10-335-977-8346	Sequence 8346, App
536	26	92.9	945	14	US-10-152-381-146	Sequence 146, App	609	25	89.3	181	15	US-10-425-114-68180	Sequence 68180, A
537	26	92.9	945	14	US-10-152-400-146	Sequence 146, App	610	25	89.3	189	15	US-10-335-977-8347	Sequence 8347, App
538	26	92.9	945	14	US-10-153-585-146	Sequence 146, App	611	25	89.3	199	15	US-10-282-122A-53754	Sequence 53754, A
539	26	92.9	945	14	US-10-157-780-146	Sequence 146, App	612	25	89.3	225	15	US-10-282-122A-71450	Sequence 71450, A
540	26	92.9	945	14	US-10-157-800-146	Sequence 146, App	613	25	89.3	243	15	US-10-424-599-190595	Sequence 190595,
541	26	92.9	945	14	US-10-157-801-146	Sequence 146, App	614	25	89.3	243	16	US-10-437-963-144609	Sequence 144609,
542	26	92.9	945	14	US-10-157-802-146	Sequence 146, App	615	25	89.3	244	10	US-09-769-744A-50	Sequence 50, Appl
543	26	92.9	945	14	US-10-158-784-146	Sequence 146, App	616	25	89.3	271	15	US-10-282-122A-71810	Sequence 71810, A
544	26	92.9	945	14	US-10-158-789-146	Sequence 146, App	617	25	89.3	276	15	US-10-425-114-40354	Sequence 40354, A
545	26	92.9	945	14	US-10-192-011-146	Sequence 146, App	618	25	89.3	283	15	US-10-425-114-43426	Sequence 43426, A
546	26	92.9	945	14	US-10-139-963-146	Sequence 146, App	619	25	89.3	300	14	US-10-369-493-1306	Sequence 1306, App
547	26	92.9	945	14	US-10-140-020-146	Sequence 146, App	620	25	89.3	300	14	US-10-369-493-20371	Sequence 20371, A
548	26	92.9	945	14	US-10-140-023-146	Sequence 146, App	621	25	89.3	300	14	US-10-369-493-21594	Sequence 21594, A
549	26	92.9	945	14	US-10-140-809-146	Sequence 146, App	622	25	89.3	311	10	US-09-988-626-230	Sequence 230, App
550	26	92.9	945	14	US-10-140-865-146	Sequence 146, App	623	25	89.3	311	10	US-09-988-687-230	Sequence 230, App
551	26	92.9	945	14	US-10-141-701-146	Sequence 146, App	624	25	89.3	311	10	US-09-988-686-230	Sequence 230, App
552	26	92.9	945	14	US-10-141-754-146	Sequence 146, App	625	25	89.3	311	14	US-10-190-279-20	Sequence 20, Appl
553	26	92.9	945	14	US-10-141-760-146	Sequence 146, App	626	25	89.3	335	15	US-10-335-977-9349	Sequence 9349, App
554	26	92.9	945	14	US-10-142-425-146	Sequence 146, App	627	25	89.3	360	14	US-10-369-493-18284	Sequence 18284, A
555	26	92.9	945	14	US-10-142-430-146	Sequence 146, App	628	25	89.3	417	16	US-10-437-963-167953	Sequence 167953,
556	26	92.9	945	14	US-10-143-113-146	Sequence 146, App	629	25	89.3	432	15	US-10-282-122A-54524	Sequence 54524, A
557	26	92.9	945	14	US-10-146-730-146	Sequence 146, App	630	25	89.3	442	15	US-10-282-122A-52544	Sequence 52544, A
558	26	92.9	945	14	US-10-146-792-146	Sequence 146, App	631	25	89.3	484	13	US-10-087-192-1074	Sequence 1074, App
559	26	92.9	945	14	US-10-158-791-146	Sequence 146, App	632	25	89.3	485	16	US-10-437-963-118960	Sequence 118960,
560	26	92.9	945	14	US-10-156-843-146	Sequence 146, App	633	25	89.3	509	14	US-10-156-761-8929	Sequence 8929, App
561	26	92.9	945	14	US-10-157-786-146	Sequence 146, App	634	25	89.3	583	14	US-10-369-493-4893	Sequence 4893, App
562	26	92.9	945	14	US-10-152-405-146	Sequence 146, App	635	25	89.3	586	16	US-10-437-963-130269	Sequence 130269,
563	26	92.9	945	14	US-10-147-528-146	Sequence 146, App	636	25	89.3	681	15	US-10-282-122A-73072	Sequence 73072, A
564	26	92.9	945	15	US-10-128-692A-146	Sequence 146, App	637	25	89.3	683	15	US-10-335-977-9350	Sequence 9350, App
565	26	92.9	945	15	US-10-140-927-146	Sequence 146, App	638	25	89.3	845	16	US-10-437-963-149418	Sequence 149418,
566	26	92.9	945	15	US-10-087-684-33	Sequence 33, Appl	639	25	89.3	852	16	US-10-437-963-118390	Sequence 118390,
567	26	92.9	945	15	US-10-087-684-34	Sequence 34, Appl	640	25	89.3	852	9	US-09-815-242-10151	Sequence 10151, A
568	26	92.9	945	15	US-10-147-493-146	Sequence 146, App	641	25	89.3	865	15	US-10-282-122A-56539	Sequence 56539, A
569	26	92.9	945	15	US-10-218-779-33	Sequence 33, Appl	642	25	89.3	865	15	US-10-282-122A-59478	Sequence 59478, A
570	26	92.9	945	15	US-10-145-127-146	Sequence 146, App	643	25	89.3	865	15	US-10-282-122A-59478	Sequence 59478, A
571	26	92.9	945	15	US-10-160-503-146	Sequence 146, App	644	25	89.3	865	15	US-10-282-122A-75994	Sequence 75994, A
572	26	92.9	945	15	US-10-143-118-146	Sequence 146, App	645	25	89.3	867	9	US-09-815-242-13848	Sequence 13848, A
573	26	92.9	945	15	US-10-143-118-146	Sequence 146, App	646	25	89.3	867	9	US-09-815-242-5122	Sequence 5122, App
574	26	92.9	945	15	US-10-158-787-146	Sequence 146, App	647	25	89.3	868	9	US-09-815-242-11233	Sequence 11233, A
575	26	92.9	945	15	US-10-037-417-38	Sequence 38, Appl	648	25	89.3	868	15	US-10-282-122A-43448	Sequence 43448, A
576	26	92.9	945	15	US-10-037-417-121	Sequence 121, App	649	25	89.3	868	15	US-10-282-122A-58502	Sequence 58502, A
577	26	92.9	945	15	US-10-140-024-146	Sequence 146, App	650	25	89.3	869	15	US-10-282-122A-67602	Sequence 67602, A
578	26	92.9	945	15	US-10-096-625-123	Sequence 123, App	651	25	89.3	869	15	US-10-282-122A-69531	Sequence 69531, A
579	26	92.9	945	15	US-10-096-625-124	Sequence 124, App	652	25	89.3	870	14	US-10-369-493-13891	Sequence 13891, A
580	26	92.9	945	15	US-10-147-536-146	Sequence 146, App	653	25	89.3	912	9	US-09-865-960-2	Sequence 2, Appl
581	26	92.9	945	16	US-10-152-372-146	Sequence 146, App	654	25	89.3	912	15	US-10-448-162-2	Sequence 162935,
582	26	92.9	945	17	US-10-788-792-252	Sequence 252, App	655	25	89.3	1034	16	US-10-437-963-182935	Sequence 182935,
583	26	92.9	945	17	US-10-125-795-146	Sequence 146, App	656	25	89.3	1326	14	US-10-369-493-22887	Sequence 22887, A
584	26	92.9	946	11	US-09-972-211-56	Sequence 56, Appl	657	25	89.3	1328	14	US-10-369-493-18867	Sequence 18867, A
585	26	92.9	946	11	US-10-096-625-56	Sequence 56, Appl	658	25	89.3	1331	14	US-10-369-493-20000	Sequence 20000, A
586	26	92.9	946	11	US-10-096-625-56	Sequence 56, Appl	659	25	89.3	1331	14	US-10-369-493-2641	Sequence 2641, App
587	26	92.9	950	14	US-10-295-027-1340	Sequence 1340, App	660	25	89.3	1349	14	US-10-369-493-20902	Sequence 20902, A
588	26	92.9	1015	14	US-10-369-493-6805	Sequence 6805, App	661	25	89.3	1587	16	US-10-437-963-116498	Sequence 116498,
589	26	92.9	1015	14	US-10-156-761-9633	Sequence 9633, App	662	25	89.3	1658	16	US-10-437-963-155728	Sequence 155728,
590	26	92.9	1058	15	US-10-282-122A-74498	Sequence 74498, A	663	25	89.3	1736	16	US-10-437-963-190641	Sequence 190641,
591	26	92.9	1059	15	US-10-282-122A-72291	Sequence 72291, A	664	25	89.3	1788	16	US-10-437-963-143965	Sequence 143965,
592	26	92.9	1060	15	US-10-282-122A-57075	Sequence 57075, A	665	25	89.3	1811	16	US-10-437-963-195747	Sequence 195747,
593	26	92.9	1061	15	US-10-282-122A-57652	Sequence 57652, A	666	25	89.3	2003	16	US-10-437-963-119004	Sequence 119004,
594	26	92.9	1062	14	US-10-369-493-17331	Sequence 17331, App	667	25	89.3	2068	16	US-10-437-963-195562	Sequence 195562,
595	26	92.9	1088	16	US-10-408-765A-1969	Sequence 1969, App	668	24	85.7	13	10	US-09-789-996-16	Sequence 16, Appl
596	26	92.9	1329	14	US-10-369-493-1221	Sequence 1221, App	669	24	85.7	68	15	US-10-424-599-171936	Sequence 171936,
597	26	92.9	1384	15	US-10-424-599-192608	Sequence 192608,	670	24	85.7	95	16	US-10-437-963-106252	Sequence 106252,

671	24	85.7	98	15	US-10-335-977-6165	Sequence 6165, Ap	744	23	82.1	51	15	US-10-424-599-258112	Sequence 258112,
672	24	85.7	143	15	US-10-382-122A-70716	Sequence 70716, A	745	23	82.1	51	16	US-10-437-963-147144	Sequence 147144,
673	24	85.7	143	15	US-10-382-122A-71638	Sequence 71638, A	746	23	82.1	51	15	US-10-424-599-249850	Sequence 249850,
674	24	85.7	152	16	US-10-437-963-156431	Sequence 156431, A	747	23	82.1	53	15	US-10-424-599-272287	Sequence 272287,
675	24	85.7	159	16	US-10-767-701-62354	Sequence 62354, A	748	23	82.1	61	15	US-10-424-598-266360	Sequence 266360,
676	24	85.7	201	9	US-09-738-626-6916	Sequence 6916, Ap	749	23	82.1	62	15	US-10-424-599-285070	Sequence 285070,
677	24	85.7	206	16	US-10-781-014-608	Sequence 608, Ap	750	23	82.1	68	16	US-10-302-547-61	Sequence 61, Appl
678	24	85.7	218	15	US-10-425-114-59362	Sequence 59362, A	751	23	82.1	73	9	US-09-993-844-63	Sequence 63, Appl
679	24	85.7	223	16	US-10-767-701-39554	Sequence 39554, A	752	23	82.1	73	15	US-10-424-599-275538	Sequence 275538,
680	24	85.7	224	9	US-09-942-935-2	Sequence 2, Appli	753	23	82.1	76	16	US-10-437-963-118997	Sequence 118997,
681	24	85.7	255	16	US-10-437-963-187525	Sequence 187525,	754	23	82.1	78	14	US-10-059-911-20	Sequence 20, Appl
682	24	85.7	268	14	US-10-156-761-14788	Sequence 14788, A	755	23	82.1	79	15	US-10-424-599-229797	Sequence 229797,
683	24	85.7	296	14	US-10-369-493-23381	Sequence 23381, A	756	23	82.1	79	16	US-10-437-963-149206	Sequence 149206,
684	24	85.7	311	16	US-10-767-701-43841	Sequence 43841, A	757	23	82.1	80	15	US-10-424-599-209524	Sequence 209524,
685	24	85.7	315	13	US-10-040-655-9	Sequence 9, Appli	758	23	82.1	82	15	US-10-424-599-275537	Sequence 275537,
686	24	85.7	315	13	US-10-041-006A-9	Sequence 9, Appli	759	23	82.1	83	9	US-09-993-844-64	Sequence 64, Appl
687	24	85.7	350	10	US-09-789-996-23	Sequence 23, Appl	760	23	82.1	83	15	US-10-424-599-230404	Sequence 230404,
688	24	85.7	350	10	US-09-952-680A-24	Sequence 24, Appl	761	23	82.1	84	16	US-10-437-963-168404	Sequence 168404,
689	24	85.7	350	15	US-10-352-843-15	Sequence 15, Appl	762	23	82.1	86	14	US-10-038-010-16	Sequence 16, Appl
690	24	85.7	350	15	US-10-380-393B-1	Sequence 1, Appli	763	23	82.1	88	16	US-10-437-963-154185	Sequence 154185,
691	24	85.7	350	15	US-10-059-266B-8	Sequence 8, Appli	764	23	82.1	89	15	US-10-424-599-207490	Sequence 207490,
692	24	85.7	350	16	US-10-408-765A-428	Sequence 428, App	765	23	82.1	90	16	US-10-437-963-140431	Sequence 140431,
693	24	85.7	351	15	US-10-482-122A-51593	Sequence 51593, A	766	23	82.1	92	9	US-09-993-844-21	Sequence 21, Appl
694	24	85.7	358	15	US-10-382-122A-57801	Sequence 57801, A	767	23	82.1	92	15	US-10-633-438-7	Sequence 7, Appli
695	24	85.7	359	15	US-10-282-122A-57074	Sequence 57074, A	768	23	82.1	101	15	US-10-424-599-167273	Sequence 167273,
696	24	85.7	369	14	US-10-369-493-19998	Sequence 19998, A	769	23	82.1	105	15	US-10-282-122A-62494	Sequence 62494, A
697	24	85.7	370	14	US-10-369-493-18602	Sequence 18602, A	770	23	82.1	105	15	US-10-282-122A-63949	Sequence 63949, A
698	24	85.7	373	15	US-10-432-443-23	Sequence 23, Appl	771	23	82.1	105	15	US-10-282-122A-64436	Sequence 64436, A
699	24	85.7	373	15	US-10-432-443-31	Sequence 31, Appl	772	23	82.1	105	15	US-10-424-599-223426	Sequence 223426,
700	24	85.7	374	14	US-10-369-493-12584	Sequence 12584, A	773	23	82.1	110	16	US-10-059-911-18	Sequence 18, Appl
701	24	85.7	381	16	US-10-781-014-620	Sequence 620, App	774	23	82.1	110	16	US-10-437-963-181795	Sequence 181795,
702	24	85.7	383	9	US-09-738-626-4233	Sequence 4233, Ap	775	23	82.1	112	15	US-10-424-599-201188	Sequence 201188,
703	24	85.7	385	14	US-10-369-493-18107	Sequence 18107, A	776	23	82.1	112	15	US-10-424-599-264226	Sequence 264226,
704	24	85.7	386	14	US-10-369-493-13478	Sequence 13478, A	777	23	82.1	118	15	US-10-425-114-63765	Sequence 63765, A
705	24	85.7	394	9	US-09-738-626-4273	Sequence 4273, Ap	778	23	82.1	121	15	US-10-424-599-254917	Sequence 254917,
706	24	85.7	396	15	US-10-282-122A-71677	Sequence 71677, A	779	23	82.1	121	15	US-10-389-647-373	Sequence 373, App
707	24	85.7	399	15	US-10-424-599-256221	Sequence 256221, A	780	23	82.1	123	15	US-10-424-599-178294	Sequence 178294,
708	24	85.7	449	15	US-10-282-122A-48267	Sequence 48267, A	781	23	82.1	129	15	US-10-425-114-49079	Sequence 49079, A
709	24	85.7	478	15	US-10-425-114-51037	Sequence 51037, A	782	23	82.1	130	15	US-10-424-599-257582	Sequence 257582,
710	24	85.7	496	15	US-10-424-599-145305	Sequence 145305, A	783	23	82.1	130	15	US-10-424-599-267181	Sequence 267181,
711	24	85.7	496	16	US-10-437-963-174130	Sequence 174130, A	784	23	82.1	131	16	US-10-767-701-55190	Sequence 55190, A
712	24	85.7	500	16	US-10-437-963-174132	Sequence 174132, A	785	23	82.1	133	15	US-10-424-599-146245	Sequence 146245,
713	24	85.7	503	14	US-10-310-154-541	Sequence 541, App	786	23	82.1	134	15	US-10-424-599-146245	Sequence 146245,
714	24	85.7	513	9	US-09-815-242-10608	Sequence 10608, A	787	23	82.1	138	15	US-10-424-599-264748	Sequence 264748,
715	24	85.7	513	14	US-10-369-493-13018	Sequence 13018, A	788	23	82.1	138	16	US-10-437-963-159848	Sequence 159848,
716	24	85.7	552	15	US-10-282-122A-43916	Sequence 43916, A	789	23	82.1	138	16	US-10-437-963-185069	Sequence 185069,
717	24	85.7	557	14	US-10-369-493-159	Sequence 159, App	790	23	82.1	140	14	US-10-038-010-18	Sequence 18, Appl
718	24	85.7	557	15	US-10-335-977-6167	Sequence 6167, Ap	791	23	82.1	141	15	US-10-424-599-184867	Sequence 184867,
719	24	85.7	564	14	US-10-369-493-12501	Sequence 12501, A	792	23	82.1	157	11	US-09-864-408A-2644	Sequence 2644, Ap
720	24	85.7	568	15	US-10-425-114-70665	Sequence 70665, A	793	23	82.1	159	16	US-10-437-963-168769	Sequence 168769,
721	24	85.7	592	14	US-10-032-585-7396	Sequence 7396, Ap	794	23	82.1	161	13	US-10-001-873-52	Sequence 52, Appl
722	24	85.7	611	16	US-10-437-963-149539	Sequence 149539, A	795	23	82.1	165	16	US-10-437-963-148307	Sequence 148307,
723	24	85.7	637	14	US-10-369-493-16994	Sequence 16994, A	796	23	82.1	168	15	US-10-425-114-46575	Sequence 46575, A
724	24	85.7	642	15	US-10-425-114-52198	Sequence 52198, A	797	23	82.1	171	14	US-10-156-761-9002	Sequence 9002, Ap
725	24	85.7	726	15	US-10-282-122A-52202	Sequence 52202, A	798	23	82.1	171	16	US-10-437-963-159850	Sequence 159850,
726	24	85.7	731	14	US-10-369-493-12947	Sequence 12947, A	799	23	82.1	172	15	US-10-424-599-162423	Sequence 162423,
727	24	85.7	771	14	US-10-369-493-10398	Sequence 10398, A	800	23	82.1	174	14	US-10-059-911-19	Sequence 19, Appl
728	24	85.7	774	16	US-10-437-963-148764	Sequence 148764, A	801	23	82.1	177	14	US-10-369-493-173	Sequence 173, App
729	24	85.7	853	14	US-10-032-585-7830	Sequence 7830, Ap	802	23	82.1	178	14	US-10-369-493-251	Sequence 251, App
730	24	85.7	865	15	US-09-751-687-16	Sequence 16, Appl	803	23	82.1	179	15	US-10-282-122A-76617	Sequence 76617, A
731	24	85.7	866	15	US-10-424-599-239573	Sequence 239573, A	804	23	82.1	180	15	US-10-282-122A-47162	Sequence 47162, A
732	24	85.7	878	14	US-10-369-493-14441	Sequence 14441, A	805	23	82.1	185	15	US-10-425-114-57524	Sequence 57524, A
733	24	85.7	1035	16	US-10-437-963-126560	Sequence 126560, A	806	23	82.1	187	16	US-10-437-963-120961	Sequence 120961,
734	24	85.7	1040	15	US-10-282-122A-48897	Sequence 48897, A	807	23	82.1	188	15	US-10-424-599-254545	Sequence 254545,
735	24	85.7	1046	15	US-10-282-122A-59531	Sequence 59531, A	808	23	82.1	192	14	US-10-038-010-20	Sequence 20, Appl
736	24	85.7	1071	14	US-10-369-493-23117	Sequence 23117, A	809	23	82.1	193	15	US-10-425-114-46437	Sequence 46437, A
737	24	85.7	1104	14	US-10-369-493-11460	Sequence 11460, A	810	23	82.1	195	15	US-10-424-599-175252	Sequence 175252,
738	24	85.7	1104	14	US-10-369-493-14726	Sequence 14726, A	811	23	82.1	196	15	US-10-425-114-51032	Sequence 51032, A
739	24	85.7	1104	14	US-10-369-493-15207	Sequence 15207, A	812	23	82.1	197	15	US-10-424-599-220042	Sequence 220042,
740	24	85.7	1138	16	US-10-437-963-152562	Sequence 152562, A	813	23	82.1	203	15	US-10-425-114-60903	Sequence 60903, A
741	24	85.7	1972	15	US-10-276-312-2	Sequence 2, Appli	814	23	82.1	204	15	US-10-424-599-258540	Sequence 258540,
742	23	82.1	18	14	US-10-225-567A-814	Sequence 814, App	815	23	82.1	205	14	US-10-059-911-17	Sequence 17, Appl
743	23	82.1	38	15	US-10-424-599-263924	Sequence 263924, A	816	23	82.1	207	14	US-10-369-493-13297	Sequence 13297, A

817	23	82.1	207	15	US-10-282-1222A-54043	Sequence 54043, A	890	23	82.1	329	15	US-10-425-114-48264	Sequence 48264, A
818	23	82.1	209	14	US-10-225-066A-404	Sequence 404, App	891	23	82.1	331	15	US-10-282-1222A-77683	Sequence 77683, A
819	23	82.1	209	14	US-10-374-780A-2762	Sequence 2762, App	892	23	82.1	332	15	US-10-425-114-58093	Sequence 58093, A
820	23	82.1	211	14	US-09-921-650-34	Sequence 34, Appl	893	23	82.1	333	15	US-10-425-114-58093	Sequence 58093, A
821	23	82.1	211	14	US-10-326-671-42	Sequence 42, Appl	894	23	82.1	333	16	US-10-425-114-58093	Sequence 58093, A
822	23	82.1	212	15	US-10-282-1222A-68295	Sequence 68295, A	895	23	82.1	334	9	US-09-159-469-28	Sequence 28, Appl
823	23	82.1	215	14	US-10-326-671-69	Sequence 69, Appl	896	23	82.1	334	9	US-09-159-469-28	Sequence 28, Appl
824	23	82.1	216	14	US-10-156-761-10916	Sequence 10916, A	897	23	82.1	334	9	US-09-159-469-28	Sequence 28, Appl
825	23	82.1	220	14	US-10-424-599-151273	Sequence 151273, A	898	23	82.1	335	10	US-09-866-050A-186	Sequence 186, App
826	23	82.1	227	14	US-10-369-493-22847	Sequence 22847, A	899	23	82.1	337	14	US-10-369-493-22847	Sequence 18342, A
827	23	82.1	236	15	US-10-425-114-39364	Sequence 3364, A	900	23	82.1	339	15	US-10-282-1222A-52554	Sequence 52554, A
828	23	82.1	242	15	US-10-424-599-223957	Sequence 223957, A	901	23	82.1	340	15	US-10-425-114-58728	Sequence 58728, A
829	23	82.1	244	14	US-10-369-493-16506	Sequence 16506, A	902	23	82.1	341	9	US-09-815-242-5410	Sequence 5410, App
830	23	82.1	245	15	US-10-282-1222A-51884	Sequence 51884, A	903	23	82.1	341	15	US-10-282-1222A-60701	Sequence 60701, A
831	23	82.1	250	16	US-10-767-701-33550	Sequence 33550, A	904	23	82.1	345	15	US-10-282-1222A-47685	Sequence 47685, A
832	23	82.1	252	15	US-10-424-599-238729	Sequence 238729, A	905	23	82.1	346	15	US-10-425-114-40247	Sequence 40247, A
833	23	82.1	254	14	US-10-032-585-7511	Sequence 7511, App	906	23	82.1	347	15	US-10-282-1222A-72163	Sequence 72163, A
834	23	82.1	254	15	US-10-425-114-61267	Sequence 61267, A	907	23	82.1	352	9	US-09-816-028A-36	Sequence 36, Appl
835	23	82.1	256	14	US-10-369-493-1710	Sequence 1710, App	908	23	82.1	352	9	US-09-815-242-12365	Sequence 12365, A
836	23	82.1	259	15	US-10-282-1222A-52957	Sequence 52957, A	909	23	82.1	352	14	US-10-087-714-2	Sequence 2, Appl
837	23	82.1	264	9	US-09-886-319A-67	Sequence 67, Appl	910	23	82.1	352	14	US-10-303-161-36	Sequence 36, Appl
838	23	82.1	264	14	US-10-376-564-67	Sequence 67, Appl	911	23	82.1	352	14	US-10-303-118-36	Sequence 36, Appl
839	23	82.1	266	16	US-10-437-963-147773	Sequence 147773, A	912	23	82.1	352	14	US-10-303-128-36	Sequence 36, Appl
840	23	82.1	267	14	US-10-369-493-11713	Sequence 11713, A	913	23	82.1	352	14	US-10-303-134-36	Sequence 36, Appl
841	23	82.1	267	14	US-10-369-493-14227	Sequence 14227, A	914	23	82.1	352	14	US-10-303-162-36	Sequence 36, Appl
842	23	82.1	267	14	US-10-369-493-14800	Sequence 14800, A	915	23	82.1	352	15	US-10-282-1222A-43870	Sequence 43870, A
843	23	82.1	267	14	US-10-369-493-14974	Sequence 14974, A	916	23	82.1	352	15	US-10-282-1222A-70705	Sequence 70705, A
844	23	82.1	267	16	US-10-437-963-168904	Sequence 168904, A	917	23	82.1	352	15	US-10-282-1222A-71375	Sequence 71375, A
845	23	82.1	270	15	US-10-282-1222A-67194	Sequence 67194, A	918	23	82.1	352	17	US-10-820-536-36	Sequence 36, Appl
846	23	82.1	271	15	US-10-389-566-1769	Sequence 1769, App	919	23	82.1	352	17	US-10-845-408-36	Sequence 36, Appl
847	23	82.1	275	14	US-10-059-911-16	Sequence 16, Appl	920	23	82.1	352	17	US-10-845-412-36	Sequence 36, Appl
848	23	82.1	275	14	US-10-059-911-25	Sequence 25, Appl	921	23	82.1	353	15	US-10-424-599-195993	Sequence 195993, A
849	23	82.1	277	14	US-10-059-911-11	Sequence 11, Appl	922	23	82.1	353	15	US-10-425-114-44266	Sequence 44266, A
850	23	82.1	282	9	US-09-764-864-1174	Sequence 1174, App	923	23	82.1	353	16	US-10-437-963-178169	Sequence 178169, A
851	23	82.1	287	15	US-10-425-114-36754	Sequence 36754, A	924	23	82.1	354	15	US-10-389-566-1765	Sequence 1765, App
852	23	82.1	289	14	US-10-369-493-9776	Sequence 9776, App	925	23	82.1	355	9	US-09-895-913A-240	Sequence 240, App
853	23	82.1	290	15	US-10-425-114-60883	Sequence 60883, A	926	23	82.1	355	14	US-10-369-493-22901	Sequence 22901, A
854	23	82.1	290	16	US-10-363-829-289	Sequence 289, App	927	23	82.1	357	14	US-10-059-911-14	Sequence 14, Appl
855	23	82.1	292	15	US-10-282-1222A-65839	Sequence 65839, A	928	23	82.1	358	14	US-10-291-172-189	Sequence 189, App
856	23	82.1	295	15	US-10-424-599-272882	Sequence 272882, A	929	23	82.1	358	15	US-10-221-278-189	Sequence 189, App
857	23	82.1	295	16	US-10-437-963-122274	Sequence 122274, A	930	23	82.1	359	14	US-10-262-313-13	Sequence 13, Appl
858	23	82.1	298	15	US-10-282-1222A-49272	Sequence 4272, A	931	23	82.1	359	16	US-10-768-878-13	Sequence 13, Appl
859	23	82.1	298	16	US-10-655-799-28	Sequence 28, Appl	932	23	82.1	360	14	US-10-087-714-20	Sequence 20, Appl
860	23	82.1	299	16	US-10-767-701-44989	Sequence 44989, A	933	23	82.1	360	14	US-10-087-714-22	Sequence 22, Appl
861	23	82.1	302	9	US-09-883-720-14	Sequence 14, Appl	934	23	82.1	362	10	US-10-087-714-23	Sequence 23, Appl
862	23	82.1	302	14	US-10-274-694-6	Sequence 6, Appl	935	23	82.1	362	10	US-09-992-331-13	Sequence 13, Appl
863	23	82.1	305	9	US-09-883-720-12	Sequence 12, Appl	936	23	82.1	362	10	US-09-964-923A-24	Sequence 24, Appl
864	23	82.1	306	14	US-10-156-761-8651	Sequence 8651, App	937	23	82.1	363	14	US-10-087-714-24	Sequence 24, Appl
865	23	82.1	307	14	US-10-190-279-17	Sequence 17, Appl	938	23	82.1	363	15	US-10-282-1222A-60356	Sequence 60356, A
866	23	82.1	308	14	US-10-059-911-10	Sequence 10, Appl	939	23	82.1	364	14	US-10-369-493-6540	Sequence 6340, App
867	23	82.1	308	15	US-10-425-114-51874	Sequence 51874, A	940	23	82.1	364	15	US-10-424-599-143711	Sequence 143711, A
868	23	82.1	310	9	US-09-815-242-5683	Sequence 5683, App	941	23	82.1	366	15	US-10-425-114-36708	Sequence 36708, A
869	23	82.1	310	15	US-10-424-599-266327	Sequence 266327, A	942	23	82.1	366	15	US-10-282-1222A-62721	Sequence 62721, A
870	23	82.1	311	9	US-09-883-720-16	Sequence 16, Appl	943	23	82.1	368	15	US-10-282-1222A-63774	Sequence 63774, A
871	23	82.1	311	16	US-10-437-963-189658	Sequence 189658, A	944	23	82.1	368	15	US-10-282-1222A-64671	Sequence 64671, A
872	23	82.1	312	15	US-10-425-114-47459	Sequence 47459, A	945	23	82.1	368	15	US-10-425-114-42679	Sequence 42679, A
873	23	82.1	313	14	US-10-380-558-15	Sequence 15, Appl	946	23	82.1	368	15	US-10-425-114-56425	Sequence 56425, A
874	23	82.1	313	15	US-10-282-1222A-56659	Sequence 56659, A	947	23	82.1	368	15	US-10-425-114-59088	Sequence 59088, A
875	23	82.1	313	15	US-10-424-599-266325	Sequence 266325, A	948	23	82.1	368	15	US-10-425-114-69552	Sequence 69552, A
876	23	82.1	314	14	US-10-059-911-23	Sequence 23, Appl	949	23	82.1	370	14	US-10-059-911-15	Sequence 15, Appl
877	23	82.1	314	16	US-10-767-701-45216	Sequence 45216, A	950	23	82.1	371	15	US-10-282-1222A-53765	Sequence 53765, A
878	23	82.1	315	15	US-10-425-114-46993	Sequence 46993, A	951	23	82.1	371	15	US-10-282-1222A-61714	Sequence 61714, A
879	23	82.1	315	15	US-10-425-114-46993	Sequence 46993, A	952	23	82.1	373	15	US-10-425-114-39691	Sequence 39691, A
880	23	82.1	315	15	US-10-425-114-35526	Sequence 35526, A	953	23	82.1	376	15	US-10-424-599-198007	Sequence 198007, A
881	23	82.1	315	15	US-10-425-114-60313	Sequence 60313, A	954	23	82.1	378	14	US-10-219-220-158	Sequence 158, App
882	23	82.1	319	9	US-09-764-868-659	Sequence 659, App	955	23	82.1	378	14	US-10-424-599-156033	Sequence 156033, A
883	23	82.1	319	15	US-10-282-1222A-47606	Sequence 47606, A	956	23	82.1	385	14	US-10-291-990-20	Sequence 20, Appl
884	23	82.1	320	16	US-10-437-963-124203	Sequence 124203, A	957	23	82.1	385	15	US-10-328-916-16	Sequence 16, Appl
885	23	82.1	327	14	US-10-067-443-5	Sequence 5, Appl	958	23	82.1	387	16	US-10-408-765A-2232	Sequence 2232, App
886	23	82.1	327	15	US-10-282-1222A-64294	Sequence 64294, A	959	23	82.1	388	14	US-10-309-515-28	Sequence 28, Appl
887	23	82.1	327	15	US-10-649-273-5	Sequence 5, Appl	960	23	82.1	388	14	US-10-309-515-46	Sequence 46, Appl
888	23	82.1	327	15	US-10-651-722-5	Sequence 5, Appl	961	23	82.1	388	14	US-10-238-075-650	Sequence 650, App
889	23	82.1	328	15	US-10-282-1222A-63494	Sequence 63494, A	962	23	82.1	388	14	US-10-126-764-28	Sequence 28, Appl

963 23 82.1 388 14 US-10-126-764-46 Sequence 46, Appl  
964 23 82.1 388 15 US-10-282-122A-59163 Sequence 59163, A  
965 23 82.1 389 15 US-10-282-122A-60399 Sequence 60399, A  
966 23 82.1 392 14 US-10-156-761-8620 Sequence 8620, Ap  
967 23 82.1 394 14 US-10-369-493-22919 Sequence 22919, A  
968 23 82.1 397 16 US-10-311-954A-6 Sequence 6, Appli  
969 23 82.1 399 9 US-09-815-242-12405 Sequence 12405, A  
970 23 82.1 399 14 US-10-352-684A-54 Sequence 54, Appl  
971 23 82.1 399 15 US-10-282-122A-44282 Sequence 44282, A  
972 23 82.1 399 15 US-10-282-122A-74587 Sequence 74587, A  
973 23 82.1 400 16 US-10-437-963-203018 Sequence 203018,  
974 23 82.1 403 16 US-10-469-864-2 Sequence 2, Appli  
975 23 82.1 407 16 US-10-437-963-168956 Sequence 168956,  
976 23 82.1 408 15 US-10-282-122A-42903 Sequence 42903, A  
977 23 82.1 410 15 US-10-425-114-59054 Sequence 59054, A  
978 23 82.1 411 15 US-10-374-780A-1326 Sequence 1326, Ap  
979 23 82.1 413 9 US-09-811-286-2 Sequence 2, Appli  
980 23 82.1 413 9 US-09-951-622-12 Sequence 12, Appl  
981 23 82.1 413 9 US-09-993-844-2 Sequence 2, Appli  
982 23 82.1 413 9 US-09-990-596-1 Sequence 1, Appli  
983 23 82.1 413 9 US-09-922-494-13 Sequence 13, Appl  
984 23 82.1 413 10 US-09-952-680A-12 Sequence 12, Appl  
985 23 82.1 413 10 US-09-811-285-2 Sequence 2, Appli  
986 23 82.1 413 10 US-09-935-061-6 Sequence 6, Appli  
987 23 82.1 413 10 US-09-935-061-6 Sequence 8, Appli  
988 23 82.1 413 14 US-10-325-930-12 Sequence 12, Appl  
989 23 82.1 413 14 US-10-225-567A-52 Sequence 52, Appl  
990 23 82.1 413 14 US-10-309-515-24 Sequence 24, Appl  
991 23 82.1 413 14 US-10-291-990-32 Sequence 32, Appl  
992 23 82.1 413 14 US-10-126-764-24 Sequence 24, Appl  
993 23 82.1 413 15 US-10-380-196A-3 Sequence 3, Appli  
994 23 82.1 413 16 US-10-692-071-6 Sequence 6, Appli  
995 23 82.1 413 16 US-10-692-071-8 Sequence 8, Appli  
996 23 82.1 413 17 US-10-221-332A-2 Sequence 2, Appli  
997 23 82.1 417 10 US-09-935-061-10 Sequence 10, Appl  
998 23 82.1 417 16 US-10-692-071-10 Sequence 10, Appl  
999 23 82.1 418 15 US-10-282-122A-50358 Sequence 50358, A  
1000 23 82.1 419 14 US-10-369-493-6670 Sequence 6670, Ap

ALIGNMENTS

RESULT 1  
US-10-425-114-56304  
; Sequence 56304, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 56304  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700975650\_FLI.pap  
US-10-425-114-56304

Query Match 96.4%; Score 27; DB 15; Length 151;  
Best Local Similarity 83.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

.Qy 1 GYXVEE 6

Db || |||  
111 GYSVEE 116

RESULT 2

US-10-767-701-35832  
; Sequence 35832, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 35832  
; LENGTH: 216  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(216)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C75823\_1.pap  
US-10-767-701-35832

Query Match 96.4%; Score 27; DB 16; Length 216;  
Best Local Similarity 83.3%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

Db || |||  
112 GYAVEE 117

RESULT 3

US-10-369-493-9674  
; Sequence 9674, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 9674  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Desulfitobacterium hafniense  
US-10-369-493-9674

Query Match 96.4%; Score 27; DB 14; Length 279;  
Best Local Similarity 83.3%; Pred. No. 4e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6

Db || |||  
70 GYSVEE 75

RESULT 4

US-10-136-253-7  
; Sequence 7, Application US/10136253  
; Publication No. US20020136737A1  
; GENERAL INFORMATION:  
; APPLICANT: FRANKEL, Fred R  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACIDS COMPRISING LISTERIA DAL AND DAT GENES  
; FILE REFERENCE: 053893-5011-02  
; CURRENT APPLICATION NUMBER: US/10/136,253  
; CURRENT FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: 09/520,207  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: PCT/US98/24357  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: 08/972,902  
; PRIOR FILING DATE: 1997-11-18  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 282  
; TYPE: PRT  
; ORGANISM: Listeria monocytogenes  
US-10-136-253-7

Query Match 96.4%; Score 27; DB 13; Length 282;  
Best Local Similarity 83.3%; Pred. No. 4e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
Db 64 GYVEE 69

RESULT 5  
US-10-369-493-122  
; Sequence 122, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 122  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Aquifex aeolicus  
US-10-369-493-122

Query Match 96.4%; Score 27; DB 14; Length 313;  
Best Local Similarity 83.3%; Pred. No. 4.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
Db 246 GYSVEE 251

RESULT 6  
US-09-815-242-5059  
; Sequence 5059, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5059  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-815-242-5059

Query Match 96.4%; Score 27; DB 9; Length 344;  
Best Local Similarity 83.3%; Pred. No. 4.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
Db 322 GYAVEE 327

RESULT 7  
US-10-282-122A-43511  
; Sequence 43511, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43511
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-43511

Query Match          96.4%; Score 27; DB 15; Length 344;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 322 GYAVEE 327

RESULT 8
US-10-369-493-16467
; Sequence 16467, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16467
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-369-493-16467

Query Match          96.4%; Score 27; DB 14; Length 365;
Best Local Similarity 83.3%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 289 GYAVEE 294

RESULT 9
US-10-369-493-6541
; Sequence 6541, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16467
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-369-493-16467
```

```
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6541
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6541

Query Match          96.4%; Score 27; DB 14; Length 390;
Best Local Similarity 83.3%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 71 GYSVEE 76

RESULT 10
US-10-425-114-37084
; Sequence 37084, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37084
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3242-059-D4_FLI.psp
US-10-425-114-37084

Query Match          96.4%; Score 27; DB 15; Length 399;
Best Local Similarity 83.3%; Pred. No. 5.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 359 GYSVEE 364

RESULT 11
US-10-437-963-117783
; Sequence 117783, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
```

; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 117783  
; LENGTH: 422  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_21156C.1.pep  
US-10-437-963-117783

Query Match 96.4%; Score 27; DB 16; Length 422;  
Best Local Similarity 83.3%; Pred. No. 6.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 105 GYSVEE 110

## RESULT 12

US-10-369-493-12764  
; Sequence 12764, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 12764  
; LENGTH: 456  
; TYPE: PRT  
; ORGANISM: Aspergillus nidulans  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(456)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-12764

Query Match 96.4%; Score 27; DB 14; Length 456;  
Best Local Similarity 83.3%; Pred. No. 6.7e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 92 GYTVEE 97

## RESULT 13

US-10-354-437-32  
; Sequence 32, Application US/10354437  
; Publication No. US20040023257A1  
; GENERAL INFORMATION:  
; APPLICANT: Barton, Nelson R.  
; APPLICANT: Weiner, David Paul  
; APPLICANT: Greenberg, William  
; APPLICANT: Luu, Samantha  
; APPLICANT: Chang, Kristine  
; APPLICANT: Waters, Elizabeth  
; TITLE OF INVENTION: ENZYMES HAVING SECONDARY AMIDASES ACTIVITY  
; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
; FILE REFERENCE: 09010-106001  
; CURRENT APPLICATION NUMBER: US/10/354,437  
; CURRENT FILING DATE: 2003-01-20  
; PRIOR APPLICATION NUMBER: US 60/352,895  
; PRIOR FILING DATE: 2002-01-28

; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 32  
; LENGTH: 474  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Obtained from an environmental sample  
US-10-354-437-32

Query Match 96.4%; Score 27; DB 15; Length 474;  
Best Local Similarity 83.3%; Pred. No. 6.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 293 GYAVEE 298

## RESULT 14

US-10-425-114-55451  
; Sequence 55451, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 55451  
; LENGTH: 508  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY077B08\_FLI.pep  
US-10-425-114-55451

Query Match 96.4%; Score 27; DB 15; Length 508;  
Best Local Similarity 83.3%; Pred. No. 7.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 468 GYSVEE 473

## RESULT 15

US-10-424-599-220386  
; Sequence 220386, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 220386  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT MRT3847\_41039C.1.pbp  
US-10-424-599-220386

Query Match 96.4%; Score 27; DB 15; Length 524;  
Best Local Similarity 83.3%; Pred. No. 7.7e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 465 GYSVEE 470

## RESULT 16

US-09-815-242-5591  
; Sequence 5591, Application US/09815242  
; Patent No. US20020061569A1

## GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5591  
; LENGTH: 545  
; TYPE: PRT

; ORGANISM: Staphylococcus aureus  
US-09-815-242-5591

Query Match 96.4%; Score 27; DB 9; Length 545;  
Best Local Similarity 83.3%; Pred. No. 8.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 269 GYSVEE 274

## RESULT 17

US-09-815-242-12198

; Sequence 12198, Application US/09815242  
; Patent No. US20020061569A1

## GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12198  
; LENGTH: 553  
; TYPE: PRT

; ORGANISM: Staphylococcus aureus  
US-09-815-242-12198

Query Match 96.4%; Score 27; DB 9; Length 553;  
Best Local Similarity 83.3%; Pred. No. 8.2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6  
|||  
Db 270 GYSVEE 275

## RESULT 18

US-10-282-122A-44061  
; Sequence 44061, Application US/10282122A  
; Publication No. US20040029129A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931



; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 44061  
; LENGTH: 553  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-282-122A-44061

Query Match 96.4%; Score 27; DB 15; Length 553;  
Best Local Similarity 83.3%; Pred. No. 8.2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 270 GYSVEE 275

## RESULT 19

US-10-437-963-173223  
; Sequence 173223, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 173223  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_71282C.1.pap  
US-10-437-963-173223

Query Match 96.4%; Score 27; DB 16; Length 575;  
Best Local Similarity 83.3%; Pred. No. 8.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 499 GYTVEE 504

## RESULT 20

US-10-389-566-1364  
; Sequence 1364, Application US/10389566  
; Publication No. US20040025202A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology, LLC  
; APPLICANT: Laurie, Cathy C  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
; FILE REFERENCE: 38-77(52900)D  
; CURRENT APPLICATION NUMBER: US/10/389,566  
; CURRENT FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: US 60/365,301  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/391,786

; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2459  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1364  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Neurospora crassa  
US-10-389-566-1364

Query Match 96.4%; Score 27; DB 15; Length 653;  
Best Local Similarity 83.3%; Pred. No. 9.8e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 492 GYSVEE 497

## RESULT 21

US-10-369-493-9757  
; Sequence 9757, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 9757  
; LENGTH: 717  
; TYPE: PRT  
; ORGANISM: Desulfitobacterium hafniense  
US-10-369-493-9757

Query Match 96.4%; Score 27; DB 14; Length 717;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 522 GYTVEE 527

## RESULT 22

US-10-369-493-18955  
; Sequence 18955, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 18955  
; LENGTH: 742

; TYPE: PRT  
; ORGANISM: Anabaena PCC7120  
US-10-369-493-18955

Query Match 96.4%; Score 27; DB 14; Length 742;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 67 GYTVEE 72

## RESULT 23

US-09-391-340-10  
; Sequence 10, Application US/09391340A  
; Patent No. US20020013455A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Mather, Eric  
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES  
; FILE REFERENCE: 09010/027001  
; CURRENT APPLICATION NUMBER: US/09/391,340A  
; CURRENT FILING DATE: 1999-09-07  
; EARLIER APPLICATION NUMBER: US 08/907,166  
; EARLIER FILING DATE: 1997-08-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10  
; LENGTH: 762  
; TYPE: PRT  
; ORGANISM: Desulfurococcus sp.  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (601)..(601)  
; OTHER INFORMATION: Xaa at position 601 is alanine or proline  
US-09-391-340-10

Query Match 96.4%; Score 27; DB 9; Length 762;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 689 GYSVEE 694

## RESULT 24

US-09-948-369-10  
; Sequence 10, Application US/09948369  
; Patent No. US20020132243A1  
; GENERAL INFORMATION:  
; APPLICANT: DIVERSA CORPORATION  
; APPLICANT: CALLEN, Walter  
; APPLICANT: MATHER, Eric  
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS C  
; FILE REFERENCE: DIVER1350-3  
; CURRENT APPLICATION NUMBER: US/09/948,369  
; CURRENT FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: US 09/656,309  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: US 09/391,340  
; PRIOR FILING DATE: 1999-09-07  
; PRIOR APPLICATION NUMBER: US 08/907,166  
; PRIOR FILING DATE: 1997-08-06  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 762  
; TYPE: PRT  
; ORGANISM: Desulfurococcus sp.  
; FEATURE:

; NAME/KEY: VARIANT  
; LOCATION: (601)..(601)  
; OTHER INFORMATION: Xaa is alanine or proline  
US-09-948-369-10

Query Match 96.4%; Score 27; DB 9; Length 762;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 689 GYSVEE 694

## RESULT 25

US-10-416-330-24  
; Sequence 24, Application US/10416330  
; Publication No. US2004019782A1  
; GENERAL INFORMATION:  
; APPLICANT: EICHMULLER, STEFAN  
; APPLICANT: SCHADENDORF, DIRK  
; APPLICANT: USENER, DIRK  
; TITLE OF INVENTION: NOVEL MARKER FOR THE DIAGNOSIS AND THERAPY OF TUMORS  
; FILE REFERENCE: 39485-0014  
; CURRENT APPLICATION NUMBER: US/10/416,330  
; CURRENT FILING DATE: 2003-05-08  
; PRIOR APPLICATION NUMBER: PCT/DE01/04229  
; PRIOR FILING DATE: 2001-11-08  
; PRIOR APPLICATION NUMBER: DE 10055285.4  
; PRIOR FILING DATE: 2000-11-08  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 1075  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-416-330-24

Query Match 96.4%; Score 27; DB 17; Length 1075;  
Best Local Similarity 83.3%; Pred. No. 1.7e+03;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 13 GYSVEE 18

## RESULT 26

US-10-282-122A-46029  
; Sequence 46029, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46029  
; LENGTH: 1405  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-10-282-122A-46029

Query Match 96.4%; Score 27; DB 15; Length 1405;  
Best Local Similarity 83.3%; Pred. No. 2.2e+03;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVEE 6  
|||  
Db 285 GYAVEE 290

## RESULT 27

US-10-424-599-237501  
; Sequence 237501, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 237501  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(59)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_56489C.1.pep  
US-10-424-599-237501

Query Match 92.9%; Score 26; DB 15; Length 59;  
Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVEE 6  
|||  
Db 35 GYGVVE 40

## RESULT 28

US-10-424-599-212021  
; Sequence 212021, Application US/10424599  
; Publication No. US20040031072A1

; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 212021  
; LENGTH: 67  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_33481C.1.pep  
US-10-424-599-212021

Query Match 92.9%; Score 26; DB 15; Length 67;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GYXVEE 6  
|||  
Db 36 GYRVEE 41

## RESULT 29

US-09-864-761-41934  
; Sequence 41934, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,455  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687

```
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41934
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL13244.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.7
; OTHER INFORMATION: SWISSPROT HIT: Q10175, EVALUE 1.20e-01
; OTHER INFORMATION: EST_HUMAN HIT: A1183641.1, EVALUE 3.00e-26
US-09-864-761-41934

Query Match          92.9%; Score 26; DB 9; Length 84;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      13 GYLVEE 18

RESULT 30
US-10-767-701-40109
; Sequence 40109, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40109
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C43438_1.pap
US-10-767-701-40109

Query Match          92.9%; Score 26; DB 16; Length 87;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      57 GYRVEE 62

RESULT 31
US-09-864-761-42006
; Sequence 42006, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
```

```
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42006
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121767.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
; OTHER INFORMATION: EST_HUMAN HIT: BF058612.1, EVALUE 1.00e-35
; OTHER INFORMATION: SWISSPROT HIT: P29539, EVALUE 6.60e-01
US-09-864-761-42006

Query Match          92.9%; Score 26; DB 9; Length 90;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      49 GYLVEE 54

RESULT 32
US-10-335-977-8299
; Sequence 8299, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
```

```
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 8299:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...95
; SEQUENCE DESCRIPTION: SEQ ID NO: 8299:
US-10-335-977-8299

Query Match 92.9%; Score 26; DB 15; Length 95;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 60 GYTIEE 65

RESULT 33
US-09-933-261-8
; Sequence 8, Application US/09933261
; Publication No. US20030040046A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030040046A1 Relevant
; TOPOLOGY: No. US20030040046A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-933-261-8

Query Match 92.9%; Score 26; DB 10; Length 102;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 84 GYLVEE 89

RESULT 34
US-10-256-702-8
; Sequence 8, Application US/10256702
; Publication No. US20030059859A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/256,702
; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
```

```
;
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030059859A1 Relevant
; TOPOLOGY: No. US20030059859A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-256-702-8

Query Match          92.9%; Score 26; DB 14; Length 102;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 84 GYLVEE 89

RESULT 35
US-10-424-599-268557
; Sequence 268557, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 268557
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_8452C.1.pgp
US-10-424-599-268557

Query Match          92.9%; Score 26; DB 15; Length 107;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 70 GYRVEE 75

RESULT 36
US-09-864-408A-3458
; Sequence 3458, Application US/09864408A
; Publication No. US2004000947A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: No. US2004000947A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3458
; LENGTH: 112
```

```
;
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-3458

Query Match          92.9%; Score 26; DB 11; Length 112;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 26 GYVEVE 31

RESULT 37
US-10-027-450-13
; Sequence 13, Application US/10027450
; Publication No. US20020102715A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/10/027,450
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/063,423
; PRIOR FILING DATE: 1997 October 28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 13
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (104)
; NAME/KEY: UNSURE
; LOCATION: (112)
US-10-027-450-13

Query Match          92.9%; Score 26; DB 13; Length 115;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 33 GYKVEE 38

RESULT 38
US-10-410-681-20
; Sequence 20, Application US/10410681
; Publication No. US20030188334A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Gutteridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: BB1370 US NA
; CURRENT APPLICATION NUMBER: US/10/410,681
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US/09/585,173B
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 118
; TYPE: PRT
```

```
; ORGANISM: Zea mays
US-10-410-661-20

Query Match      92.9%; Score 26; DB 14; Length 118;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 23 GYLVEE 28

RESULT 39
US-10-424-599-155357
; Sequence 155357, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 155357
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(118)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_111308C.1.pep
US-10-424-599-155357

Query Match      92.9%; Score 26; DB 15; Length 118;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 28 GYVVEE 33

RESULT 40
US-10-335-977-8300
; Sequence 8300, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002

; ORGANISM: Zea mays
US-10-410-661-20

Query Match      92.9%; Score 26; DB 14; Length 118;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 23 GYLVEE 28

RESULT 39
US-10-424-599-155357
; Sequence 155357, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 155357
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(118)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_111308C.1.pep
US-10-424-599-155357

Query Match      92.9%; Score 26; DB 15; Length 119;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 84 GYTIEE 89

RESULT 41
US-10-424-599-272017
; Sequence 272017, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 272017
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(123)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87650C.1.pep
US-10-424-599-272017

Query Match      92.9%; Score 26; DB 15; Length 123;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 18 GYVVEE 23

RESULT 42
US-10-424-599-159938
; Sequence 159938, Application US/10424599
```

```

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 159938
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(132)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_115443C.1.pep
US-10-424-599-159938

Query Match          92.9%; Score 26; DB 15; Length 132;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      125 GYGVEE 130

RESULT 43
US-09-864-761-34885
; Sequence 34885, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34885
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009297.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
; OTHER INFORMATION: EST HUMAN HIT: BE180565.1, EVALUE 8.00e-11
; OTHER INFORMATION: SWISSPROT HIT: P27283, EVALUE 3.80e+00
US-09-864-761-34885

Query Match          92.9%; Score 26; DB 9; Length 139;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      11 GYDVEE 16

RESULT 44
US-10-029-386-30849
; Sequence 30849, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: Aecomica-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30849
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008580.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: SWISSPROT HIT: P70742, EVALUE 1.60e+00
US-10-029-386-30849

Query Match          92.9%; Score 26; DB 14; Length 142;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      11 GYDVEE 16

```



```
Db          56 GYLVEE 61
; Sequence 198169, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 198169
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(146)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20971C.1.pep
; US-10-424-599-198169
Query Match          92.9%; Score 26; DB 15; Length 146;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          1 GYXVEE 6
          |||||
Db          30 GYGVEE 35

RESULT 48
US-10-767-701-40676
; Sequence 40676, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40676
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C79069_1.pep
; US-10-767-701-40676
Query Match          92.9%; Score 26; DB 16; Length 146;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          1 GYXVEE 6
          |||||
Db          113 GYLVEE 118

RESULT 49
US-09-925-300-1599
; Sequence 1599, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
```

---

```
Db          56 GYLVEE 61
; Sequence 169951, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 169951
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(143)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_124480C.1.pep
; US-10-424-599-169951
Query Match          92.9%; Score 26; DB 15; Length 143;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          1 GYXVEE 6
          |||||
Db          119 GYGVEE 124

RESULT 46
US-10-424-599-216365
; Sequence 216365, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216365
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37406C.1.pep
; US-10-424-599-216365
Query Match          92.9%; Score 26; DB 15; Length 145;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          1 GYXVEE 6
          |||||
Db          15 GYVEE 20

RESULT 47
US-10-424-599-198169
```

```
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1599
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (125)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (135)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1599
```

```
Query Match          92.9%; Score 26; DB 9; Length 154;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GYXVEE 6
        |||||
Db       79 GYGVEE 84
```

## RESULT 50

```
US-10-276-774-1538
; Sequence 1538, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Fang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-Q30
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1538
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-1538
```

```
Query Match          92.9%; Score 26; DB 15; Length 160;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GYXVEE 6
        |||||
Db       75 GYGVEE 80
```

## RESULT 51

```
US-10-424-599-148963
; Sequence 148963, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 148963
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105537C.1.pap
US-10-424-599-148963
```

```
Query Match          92.9%; Score 26; DB 15; Length 163;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GYXVEE 6
        |||||
Db       39 GYSIEE 44
```

## RESULT 52

```
US-10-424-599-218619
; Sequence 218619, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218619
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: unsure
; LOCATION: (1)..(170)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_39441C.1.pap
US-10-424-599-218619
```

```
Query Match          92.9%; Score 26; DB 15; Length 170;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GYXVEE 6
        |||||
Db       124 GYGVEE 129
```

## RESULT 53

```
US-10-029-386-33430
; Sequence 33430, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
```

```
; SEQ ID NO 33430
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006354.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
; OTHER INFORMATION: SWISSPROT HIT: P36006, EVALU 3.50e+00
US-10-029-386-33430

Query Match          92.9%; Score 26; DB 14; Length 173;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 77 GYLVEE 82

RESULT 54
US-10-437-963-195791
; Sequence 195791, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195791
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91705C.1.pep
US-10-437-963-195791

Query Match          92.9%; Score 26; DB 16; Length 181;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 94 GYHVEE 99

RESULT 55
US-10-424-599-150556
; Sequence 150556, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 150556

; SEQ ID NO 33430
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(196)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106977C.1.pep
US-10-424-599-150556

Query Match          92.9%; Score 26; DB 15; Length 196;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 133 GYVVEE 138

RESULT 56
US-10-424-599-203252
; Sequence 203252, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 203252
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(217)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_25561C.1.pep
US-10-424-599-203252

Query Match          92.9%; Score 26; DB 15; Length 217;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
Db 174 GYNVEE 179

RESULT 57
US-10-424-599-150557
; Sequence 150557, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 150557
; LENGTH: 225
```

```
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(225)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106978C.1.pep
US-10-424-599-150557

Query Match          92.9%; Score 26; DB 15; Length 225;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      137 GYGVVE 142

RESULT 58
US-10-425-114-40929
; Sequence 40929, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40929
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(225)
; OTHER INFORMATION: Clone ID: LIB3051-047-H7_FLI.pep
US-10-425-114-40929

Query Match          92.9%; Score 26; DB 15; Length 253;
Best Local Similarity 83.3%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      220 GYRVVE 225

RESULT 59
US-10-424-599-158869
; Sequence 158869, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 158869
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
```

---

```
; OTHER INFORMATION: Clone ID: PAT_MRT3847_114479C.1.pep
US-10-424-599-158869

Query Match          92.9%; Score 26; DB 15; Length 254;
Best Local Similarity 83.3%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      221 GYRVVE 226

RESULT 60
US-10-424-599-150562
; Sequence 150562, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 150562
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(255)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106983C.1.pep
US-10-424-599-150562

Query Match          92.9%; Score 26; DB 15; Length 255;
Best Local Similarity 83.3%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GYXVEE 6
Db      123 GYGVVE 128

RESULT 61
US-10-410-681-22
; Sequence 22, Application US/10410681
; Publication No. US20030188334A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Guttridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: BB1370 US NA
; CURRENT APPLICATION NUMBER: US/10/410,681
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US/09/585,173B
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Zea mays
US-10-410-681-22
```

```
Query Match          92.9%; Score 26; DB 14; Length 260;
Best Local Similarity 83.3%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      26 GYLVEE 31

RESULT 62
US-10-767-701-43739
; Sequence 43739, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 43739
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C16927_1.pep
US-10-767-701-43739

Query Match          92.9%; Score 26; DB 16; Length 262;
Best Local Similarity 83.3%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      185 GYKVEE 190

RESULT 63
US-10-425-114-55122
; Sequence 55122, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55122
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700753894_FLI.pep
US-10-425-114-55122

Query Match          92.9%; Score 26; DB 15; Length 263;
Best Local Similarity 83.3%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      132 GYGVEE 137

Query Match          92.9%; Score 26; DB 14; Length 260;
Best Local Similarity 83.3%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      26 GYLVEE 31

RESULT 64
US-10-424-599-150560
; Sequence 150560, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 150560
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106980C.1.pep
US-10-424-599-150560

Query Match          92.9%; Score 26; DB 15; Length 267;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      164 GYGVEE 169

RESULT 65
US-10-425-114-71978
; Sequence 71978, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71978
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY008A12_FLI.pep
US-10-425-114-71978

Query Match          92.9%; Score 26; DB 15; Length 267;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYXVEE 6
DB      136 GYGVEE 141

RESULT 66
US-10-425-114-71979
; Sequence 71979, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71979
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY014D08_FLI.pep
US-10-425-114-71979

Query Match          92.9%; Score 26; DB 15; Length 267;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 136 GYGVVE 141

RESULT 67
US-10-425-114-52770
; Sequence 52770, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52770
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3107-018-D7_FLI.pep
US-10-425-114-52770

Query Match          92.9%; Score 26; DB 15; Length 271;
Best Local Similarity 83.3%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 140 GYGVVE 145

RESULT 68
US-10-424-599-271699
; Sequence 271699, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 271699
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87363C.1.pep
US-10-424-599-271699

Query Match          92.9%; Score 26; DB 15; Length 278;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 116 GYVVEE 121

RESULT 69
US-10-029-386-33440
; Sequence 33440, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33440
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL163303.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.91
; OTHER INFORMATION: SWISSPROT HIT: P31626, EVALUE 6.30e-01
US-10-029-386-33440

Query Match          92.9%; Score 26; DB 14; Length 286;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6
Db 129 GYLVEE 134

RESULT 70
US-10-425-114-40821
; Sequence 40821, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
```

```
; SEQ ID NO 40821
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3039-015-B9_FLI.pap
US-10-425-114-40821

Query Match      92.9%; Score 26; DB 15; Length 289;
Best Local Similarity 83.3%; Pred. No. 7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 158 GYVEE 163

RESULT 71
US-10-156-761-11627
; Sequence 11627, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11627
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11627

Query Match      92.9%; Score 26; DB 14; Length 303;
Best Local Similarity 66.7%; Pred. No. 7.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 174 GYSIEE 179

RESULT 72
US-10-027-450-15
; Sequence 15, Application US/10027450
; Publication No. US20020102715A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/10/027,450
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/063,423
; PRIOR FILING DATE: 1997 October 28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 15
; LENGTH: 307

; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-027-450-15

Query Match      92.9%; Score 26; DB 13; Length 307;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 230 GYQVEE 235

RESULT 73
US-10-425-114-37774
; Sequence 37774, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37774
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3069-022-E10_FLI.pap
US-10-425-114-37774

Query Match      92.9%; Score 26; DB 15; Length 319;
Best Local Similarity 83.3%; Pred. No. 7.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYXVEE 6
DB 286 GYRVEE 291

RESULT 74
US-09-988-626-231
; Sequence 231, Application US/09988626
; Publication No. US20030044959A1
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/988,626
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 231
; LENGTH: 326
; TYPE: PRT
```

; ORGANISM: Synechocystis sp.  
US-09-988-626-231

Query Match 92.9%; Score 26; DB 10; Length 326;  
Best Local Similarity 83.3%; Pred.No. 7.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 146 GYRVEE 151

## RESULT 75

US-09-988-687-231  
; Sequence 231, Application US/09988687  
; Publication No. US20030045704A1  
; GENERAL INFORMATION:  
; APPLICANT: Tavtigian, Sean V.  
; APPLICANT: Teng, David H.F.  
; APPLICANT: Simard, Jacques  
; APPLICANT: Rommens, Johanna M.  
; APPLICANT: Myriad Genetics, Inc.  
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility  
; FILE REFERENCE: 2318-258  
; CURRENT APPLICATION NUMBER: US/09/988,687  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 09/564,805  
; PRIOR FILING DATE: 2000-05-05  
; PRIOR APPLICATION NUMBER: US 60/107,468  
; PRIOR FILING DATE: 1998-11-06  
; PRIOR APPLICATION NUMBER: 09/434,382  
; PRIOR FILING DATE: 1999-11-05  
; NUMBER OF SEQ ID NOS: 240  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 231  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Synechocystis sp.  
US-09-988-687-231

Query Match 92.9%; Score 26; DB 10; Length 326;  
Best Local Similarity 83.3%; Pred.No. 7.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYXVEE 6  
|||  
Db 146 GYRVEE 151

Search completed: November 1, 2004, 21:50:31  
Job time : 106 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 20:58:17 ; Search time 159 Seconds  
(without alignment)  
1290.523 Million cell updates/sec

Title: US-10-030-194A-4  
Perfect score: 2950  
Sequence: 1 MKRDLHQFGNHTGTSAGS.....LSWTRPLITTSANKLSAVH 572

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2950	100.0	572	4	AAB31884 Amino aci
2	2946	99.9	572	4	AAB31883 Amino aci
3	2408.5	81.6	587	4	AAB31892 Arabidops
4	2408.5	81.6	587	4	AAB31892 Arabidops
5	2408.5	81.6	587	8	AAB31892 Arabidops
6	2291.5	77.7	587	3	AAB31892 Arabidops
7	2111	71.6	533	3	AAB31892 Arabidops
8	2110.5	71.5	532	2	AAB31892 Arabidops
9	2106.5	71.4	532	4	AAB31892 Arabidops
10	2106.5	71.4	532	4	AAB31892 Arabidops
11	2106.5	71.4	532	7	AAB31892 Arabidops
12	2106.5	71.4	532	7	AAB31892 Arabidops
13	2106.5	71.4	532	8	AAB31892 Arabidops
14	2106.5	71.4	532	8	AAB31892 Arabidops
15	2104	71.3	518	3	AAB31892 Arabidops
16	2100	71.2	531	3	AAB31892 Arabidops
17	2099	71.2	517	3	AAB31892 Arabidops
18	1802	54.3	630	2	AAB31892 Arabidops
19	1594	54.0	630	8	AAB31892 Arabidops
20	1592	54.0	625	7	AAB31892 Arabidops
21	1592	54.0	625	7	AAB31892 Arabidops
22	1592	54.0	625	8	AAB31892 Arabidops
23	1568	53.2	623	2	AAB31892 Arabidops
24	1551.5	52.6	609	8	AAB31892 Arabidops
25	1535.5	52.1	511	8	AAB31892 Arabidops

26	1325	44.9	425	2	AA020539	Protein e
27	1222.5	41.4	277	2	AAW38193	Arabidops
28	1089.5	36.9	262	2	AAW38194	Arabidops
29	940	31.9	298	8	ADL33453	Lolium pe
30	877	29.7	282	2	AAW30795	Arabidops
31	846	28.7	259	2	AAW30794	Arabidops
32	653	22.1	221	2	AAW30793	Arabidops
33	597	20.2	542	8	ADM48150	Polypepti
34	587.5	19.9	490	8	ADO62007	Transcrip
35	586	19.9	482	7	ADD30188	Plant yie
36	586	19.9	482	7	ADE37171	Plant yie
37	586	19.9	482	8	ADI41727	Plant tra
38	585.5	19.8	668	3	AAW28595	Maize Sca
39	584	19.8	653	2	AAW38178	Arabidops
40	584	19.8	653	3	AAW28569	Arabidops
41	581	19.7	427	7	ADC47020	Tomato Ls
42	581	19.7	427	7	ADC79244	Tomato Ls
43	581	19.7	428	2	AAW81753	Tomato Ls
44	581	19.7	544	7	ADC47018	Rice CIGR
45	581	19.7	544	7	ADC79242	Rice CIGR

ALIGNMENTS

RESULT 1  
AAB31884  
ID AAB31884 standard; protein; 572 AA.  
XX AAB31884;  
AC AAB31884;  
XX 15-MAY-2001 (first entry).  
DT Amino acid sequence of a mutant GRAS protein.  
DE GRAS protein; BZH gene; transcription factor; gibberellin; morphogenesis;  
KW plant development; dwarf plant; crucifer.  
XX Brassica napus.  
XX WO200109356-A1.  
XX 08-FEB-2001.  
XX 02-AUG-2000; 2000WO-FR002216.  
XX 02-AUG-1999; 99FR-00010023.  
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX Renard M, Delourme R, Barret P, Brunel D, Proger N, Tanguy X;  
XX WPI; 2001-182964/18.  
XX N-PSDB; AAF25481.  
XX New mutant nucleic acid encoding modified GRAS family protein, used to  
XX produce dwarf transgenic plants.  
XX Claim 4; Page 20-22; 28pp; French.  
XX The present sequence represents a mutant plant protein of the GRAS  
XX family. The mutant allele of the BZH gene contains a G1695A mutation  
XX resulting in the mutation E546K in the protein. GRAS proteins are  
XX transcription factors implicated in regulation of the response to  
XX gibberellin and thus in control of morphogenesis and plant development.  
XX The mutant GRAS protein is used to produce dwarf plants, specifically  
XX crucifers. Dwarf plants may be sown earlier (increasing nitrate  
XX accumulation without risking excessive stem growth during winter), and  
XX have better resistance to cold and lodging. They are also easier to  
XX harvest and allow for better monitoring of the crop  
XX Sequence 572 AA;  
SQ

Query Match 100.0%; Score 2950; DB 4; Length 572;  
Best Local Similarity 100.0%; Pred. No. 5.1e-260;  
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKRDLHOFQGNHGTSIAGSSTSSPAVFGKDKMMVKVEEDDELLGVLYKVRSEMAEV 60  
Db 1 MKRDLHOFQGNHGTSIAGSSTSSPAVFGKDKMMVKVEEDDELLGVLYKVRSEMAEV 60

Qy 61 ALKLEOLETMWNGAQEDGLAHLATDTVHYNPAEYLSWLDNMLTELNPAAATGSGNALNPE 120  
Db 61 ALKLEOLETMWNGAQEDGLAHLATDTVHYNPAEYLSWLDNMLTELNPAAATGSGNALNPE 120

Qy 121 INNNNNSFFTGGLKAI PGNAVCRRSNQFAFVDSNNKRLKPSSPDMSVTPSPAGV 180  
Db 121 INNNNNSFFTGGLKAI PGNAVCRRSNQFAFVDSNNKRLKPSSPDMSVTPSPAGV 180

Qy 181 IGTTVTTVTESTRPLILVDSQDNGVRLVHALMACAEAVQSSNLTAEALVKQIGFLAVSQ 240  
Db 181 IGTTVTTVTESTRPLILVDSQDNGVRLVHALMACAEAVQSSNLTAEALVKQIGFLAVSQ 240

Qy 241 AGAMRKVATYFAEALARRIYRLSPQTDHSLSDTLQMHFYETCPYLKFAHFTANQAIL 300  
Db 241 AGAMRKVATYFAEALARRIYRLSPQTDHSLSDTLQMHFYETCPYLKFAHFTANQAIL 300

Qy 301 EAFEGKKRVHVIDFSMNQGLQWALMALALREGPPSFRLTGIGPPAADNSDHLHEVGC 360  
Db 301 EAFEGKKRVHVIDFSMNQGLQWALMALALREGPPSFRLTGIGPPAADNSDHLHEVGC 360

Qy 361 KLAQLAEAIHVEFEYRGFVANSADLDSMLLELPSSETEAVAVNSVFLHKLGRGTGGIE 420  
Db 361 KLAQLAEAIHVEFEYRGFVANSADLDSMLLELPSSETEAVAVNSVFLHKLGRGTGGIE 420

Qy 421 KVFQVVKQIKPVI FTVVEQSNHNGPVFLDRFTESLHYSTLFDLSLEGAPSSQDKMSEV 480  
Db 421 KVFQVVKQIKPVI FTVVEQSNHNGPVFLDRFTESLHYSTLFDLSLEGAPSSQDKMSEV 480

Qy 481 YLQKQICNLVACEGPDPRVERHETLSQNSNRFGSGFAPAHLGSNAPKQASTLLALFNNGE 540  
Db 481 YLQKQICNLVACEGPDPRVERHETLSQNSNRFGSGFAPAHLGSNAPKQASTLLALFNNGE 540

RESULT 2  
AAB31883  
ID AAB31883 standard; protein; 572 AA.  
AC AAB31883;  
XX  
DT 15-MAY-2001 (first entry)  
DE Amino acid sequence of a wildtype GRAS protein.  
XX  
KW GRAS protein; BZH gene; transcription factor; gibberellin; morphogenesis;  
KW plant development; dwarf plant; crucifer.  
XX  
OS Brassica napus.  
XX  
PN WO200109356-A1.  
XX  
PD 08-FEB-2001.  
XX  
PF 02-AUG-2000; 2000WO-FR002216.  
XX  
PR 02-AUG-1999; 99FR-00010023.  
XX  
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX  
PI Renard M, Delourme R, Barret P, Brunel D, Froger N, Tanguy X;  
XX  
DR WPI; 2001-182964/18.

DR N-PSDB; AAF25480.  
XX New mutant nucleic acid encoding modified GRAS family protein, used to  
PT produce dwarf transgenic plants.  
XX Example 1; Page 16-17; 28pp; French.  
XX  
CC The present sequence represents a wild type plant protein of the GRAS  
CC family. The specification describes a mutant allele of the BZH gene,  
CC which contains a G1695A mutation resulting in the mutation E546K in the  
CC protein. GRAS proteins are transcription factors implicated in regulation  
CC of the response to gibberellins and thus in control of morphogenesis and  
CC plant development. The mutant GRAS protein is used to produce dwarf  
CC plants, specifically crucifers. Dwarf plants may be sown earlier  
CC (increasing nitrate accumulation without risking excessive stem growth  
CC during winter), and have better resistance to cold and lodging. They are  
CC also easier to harvest and allow for better monitoring of the crop  
XX  
SQ Sequence 572 AA;

Query Match 99.9%; Score 2946; DB 4; Length 572;  
Best Local Similarity 99.8%; Pred. No. 1.2e-259;  
Matches 571; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKRDLHOFQGNHGTSIAGSSTSSPAVFGKDKMMVKVEEDDELLGVLYKVRSEMAEV 60  
Db 1 MKRDLHOFQGNHGTSIAGSSTSSPAVFGKDKMMVKVEEDDELLGVLYKVRSEMAEV 60

Qy 61 ALKLEOLETMWNGAQEDGLAHLATDTVHYNPAEYLSWLDNMLTELNPAAATGSGNALNPE 120  
Db 61 ALKLEOLETMWNGAQEDGLAHLATDTVHYNPAEYLSWLDNMLTELNPAAATGSGNALNPE 120

Qy 121 INNNNNSFFTGGLKAI PGNAVCRRSNQFAFVDSNNKRLKPSSPDMSVTPSPAGV 180  
Db 121 INNNNNSFFTGGLKAI PGNAVCRRSNQFAFVDSNNKRLKPSSPDMSVTPSPAGV 180

Qy 181 IGTTVTTVTESTRPLILVDSQDNGVRLVHALMACAEAVQSSNLTAEALVKQIGFLAVSQ 240  
Db 181 IGTTVTTVTESTRPLILVDSQDNGVRLVHALMACAEAVQSSNLTAEALVKQIGFLAVSQ 240

Qy 241 AGAMRKVATYFAEALARRIYRLSPQTDHSLSDTLQMHFYETCPYLKFAHFTANQAIL 300  
Db 241 AGAMRKVATYFAEALARRIYRLSPQTDHSLSDTLQMHFYETCPYLKFAHFTANQAIL 300

Qy 301 EAFEGKKRVHVIDFSMNQGLQWALMALALREGPPSFRLTGIGPPAADNSDHLHEVGC 360  
Db 301 EAFEGKKRVHVIDFSMNQGLQWALMALALREGPPSFRLTGIGPPAADNSDHLHEVGC 360

Qy 361 KLAQLAEAIHVEFEYRGFVANSADLDSMLLELPSSETEAVAVNSVFLHKLGRGTGGIE 420  
Db 361 KLAQLAEAIHVEFEYRGFVANSADLDSMLLELPSSETEAVAVNSVFLHKLGRGTGGIE 420

Qy 421 KVFQVVKQIKPVI FTVVEQSNHNGPVFLDRFTESLHYSTLFDLSLEGAPSSQDKMSEV 480  
Db 421 KVFQVVKQIKPVI FTVVEQSNHNGPVFLDRFTESLHYSTLFDLSLEGAPSSQDKMSEV 480

Qy 481 YLQKQICNLVACEGPDPRVERHETLSQNSNRFGSGFAPAHLGSNAPKQASTLLALFNNGE 540  
Db 481 YLQKQICNLVACEGPDPRVERHETLSQNSNRFGSGFAPAHLGSNAPKQASTLLALFNNGE 540

Qy 541 GYRVEKNNGCLMSLWHTRPLITTSWAKLSAVH 572  
Db 541 GYRVEKNNGCLMSLWHTRPLITTSWAKLSAVH 572

RESULT 3  
AAE01892  
ID AAE01892 standard; protein; 587 AA.  
XX  
AC AAE01892;  
XX  
DT 31-JUL-2001 (first entry)  
XX

DE Arabidopsis thaliana transcription factor, G307.  
 KW Transcription factor; biochemical characteristic; controlling element;  
 KW structural characteristic; developmental characteristic; gene therapy;  
 KW agricultural biotechnology; plant trait modification.  
 XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX Key Location/Qualifiers  
 FH Domain 323..339  
 FT /label= Conserved\_domain  
 FT  
 XX WO200136597-A1.  
 XX 25-MAY-2001.  
 XX 14-NOV-2000; 2000WO-US031344.  
 XX 17-NOV-1999; 99US-0166228P.  
 PR 17-APR-2000; 2000US-0197899P.  
 PR 22-AUG-2000; 2000US-0227439P.  
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (CREE/) CREELMAN R.  
 PA (YUGG/) YU G.  
 PA (ADAM/) ADAM L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (HEAR/) HEARD J.  
 PA (SAMA/) SAMAHA R.  
 PA (PILG/) PILGRIM M.  
 PA (PINE/) PINEDA O.  
 PA (JIAN/) JIANG C.  
 XX Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;  
 PI Pilgrim M, Pineda O, Jiang C;  
 XX WPI; 2001-335999/35.  
 DR N-PSDB; AAD05776.  
 XX Nucleic acids encoding plant transcription factor polypeptides, useful  
 PT for altering the biochemical characteristics of plants e.g. corn, potato  
 PT and cotton plants.  
 XX Claim 4; Page 68-70; 127pp; English.  
 XX The present sequence is Arabidopsis thaliana transcription factor, G307.  
 CC The transcription factor is used for altering a plant's biochemical  
 CC characteristics. The transcription factor may be used to alter the  
 CC structure and developmental characteristics of plants such as soybean,  
 CC wheat, corn, potato, cotton, rice, oilseed rape, sunflower, alfalfa,  
 CC sugar cane, turf, banana, blackberry, blueberry, strawberry, raspberry,  
 CC cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes,  
 CC honey dew, lettuce, mango, melon, onion, papaya, peas, peppers,  
 CC pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon,  
 CC rosaceous fruits and/or vegetable brassicas. Transcription factors are  
 CC key controlling elements of biological pathways and altering expression  
 CC levels of 1 or more transcription factors can change entire biological  
 CC pathways in an organism. Therefore manipulating transcription factor  
 CC levels in plants offers great potential in agricultural biotechnology for  
 CC modifying a plant's traits. Transcription factor cDNA is useful in gene  
 CC therapy  
 XX SQ Sequence 587 AA;  
 Query Match 81.6%; Score 2408.5; DB 4; Length 587;  
 Best Local Similarity 81.9%; Pred. No. 1.7e-210;  
 Matches 490; Conservative 23; Mismatches 48; Indels 37; Gaps 10;  
 OY 1 MKRDLHFQFG--PNHGTGIAGSSSPAVFGKDKMMVKKEE-----DDELLGLVGLYKV 52  
 DB 1 MKRDHQQQRLSNHGTSSSSSSIS-----KDKMMVKKEEDGGGNMDELLAVGLYKV 54  
 OY 53 RSSEMAEVALKLEQETWVGNAQEDGLAHLATDTVHYNPAELYSWLDNMLTLPNPPAATT 112

Db 55 RSSEMAEVALKLEQETWVGNAQEDGLAHLATDTVHYNPAELYSWLDNMLTLPNPPA 114  
 OY 113 GSNALNPEINNNNSNF-FTGGDLKAIPGNVACRRSQFAFVDSSS-----NKLKPS 166  
 Db 115 SSNGLDVLPSPICGPPASDYDLKVIIPGNAI----YQFP-AIDSSSSNNQNKLSKCS 169  
 OY 167 SPDSMTVSPSP-----AGVIGTIVTIVT-----ESTRPLILVDSQDNGVRLVHALMACA 215  
 Db 170 SPDSMTVSTSTGTQIGGVIGTIVTIVT-----ESTRPLILVDSQDNGVRLVHALMACA 229  
 OY 216 EAVOSSNLTAEALVKQIGFLAVSOAGAMRVATYFAEALARRIYRLSPPTQIDHLSLD 275  
 Db 230 EAIQONNLTAEALVKQIGFLAVSOAGAMRVATYFAEALARRIYRLSPQNIIDHCLSD 289  
 OY 276 TLOMFYETCPYLKFAHFTANQAILEAFEGKRVHVIDFSMNOGLQWPAALQALREGG 335  
 Db 290 TLOMFYETCPYLKFAHFTANQAILEAFEGKRVHVIDFSMNOGLQWPAALQALREGG 349  
 OY 336 PPSFRLTGIPPAADNSDHLHEVGCKLAQLAEAHVFEYRGFVANSIADLADSLMLLRP 395  
 Db 350 PPTFRLTGIGPPAPDNDSDHLHEVGCKLAQLAEAHVFEYRGFVANSIADLADSLMLLRP 409  
 OY 396 SETRAVANVSFELHKLGRGTGGIEKVGKQIKPVIFTVVEQESNHNHGPVFLDRFTES 455  
 Db 410 SDTEAVANVSFELHKLGRGTGGIEKVGKQIKPVIFTVVEQESNHNHGPVFLDRFTES 469  
 OY 456 LHYYSTLFDLSLEGAPSSQDKVMSEVYLKQICNLVACBGPDRVERHETLSQWNRFGSSG 515  
 Db 470 LHYYSTLFDLSLEGAPSSQDKVMSEVYLKQICNLVACBGPDRVERHETLSQWNRFGSSG 529  
 OY 516 PPAHLGSNAFKQASTLLALFNGGEGYRVKNNKGLMLSWHTRPLITTSAAKLS-AVH 572  
 Db 530 LAPAHLGSNAFKQASTLLALFNGGEGYRVKNNKGLMLSWHTRPLITTSAAKLSAAH 587  
 RESULT 4  
 AAE02560  
 ID AAE02560 standard; protein; 587 AA.  
 AC AAE02560;  
 XX AAE02560;  
 DT 10-AUG-2001 (first entry)  
 XX A. thaliana transcription factor G308 homolog, G307.  
 DE Plant transcription factor; phenotype; sugar sensing characteristic;  
 KW transgenic plant; plant yield; growth; germination; photosynthesis;  
 KW glyoxylate metabolism; respiration; pathogen response; wounding response;  
 KW cell cycle regulation; pigmentation; flowering; senescence; physiology;  
 KW storage organ; metabolism.  
 XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX Key Location/Qualifiers  
 FH Domain 323..339  
 FT /note= "Conserved domain"  
 XX WO200135725-A1.  
 XX 25-MAY-2001.  
 XX 14-NOV-2000; 2000WO-US031414.  
 XX 17-NOV-1999; 99US-0166228P.  
 PR 17-APR-2000; 2000US-0197899P.  
 PR 22-AUG-2000; 2000US-0227439P.  
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (JIAN/) JIANG C.  
 PA (HEAR/) HEARD J.  
 PA (PINE/) PINEDA O.  
 PA (PILG/) PILGRIM M.

PA (ADAM/) ADAM L.  
PA (RIEC/) RIECHMANN J L.  
PA (YUGG/) YU G.  
PA (SAMA/) SAMAHA R.  
XX  
PI Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JL;  
PI Yu G, Samaha R;  
XX  
DR WPI; 2001-335977/35.  
DR N-PSDB; AAD06661.  
XX  
PT Nucleic acids encoding plant transcription factor polypeptides, useful  
PT for altering the sugar sensing characteristics of plants and increasing  
PT yield, e.g. corn, potato and cotton plants.  
XX  
PS Claim 4; Page 115-117; 151pp; English.  
XX  
CC The patent relates to polynucleotides encoding 35 plant transcription  
CC factors which may be used to modify phenotype associated with a plant's  
CC sugar sensing characteristics and increasing yield when their expression  
CC level is altered. Sugars are central regulatory molecules that control  
CC aspects of physiology, metabolism and development. Therefore the cDNAs  
CC and proteins of the invention are useful for modifying the growth and  
CC germination rates of plants, photosynthesis, glyoxylate metabolism,  
CC respiration, starch and sucrose synthesis and degradation, pathogen  
CC response, wounding response, cell cycle regulation, pigmentation,  
CC flowering and senescence of plants and for modifying sink-source  
CC relationships in seeds, tubers, roots, and other storage organs leading  
CC to an increase in yield. The transcription factor polynucleotides and  
CC polypeptides may be used to alter the structure and developmental  
CC characteristics of plants such as soybean, wheat, corn, potato, cotton,  
CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,  
CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,  
CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,  
CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,  
CC tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas.  
CC The present sequence is a homolog of Arabidopsis thaliana transcription  
CC factor  
XX  
SQ Sequence 587 AA;

Query Match 81.6%; Score 2408.5; DB 4; Length 587;  
Best Local Similarity 81.9%; Pred. No. 1.7e-210;  
Matches 490; Conservative 23; Mismatches 48; Indels 37; Gaps 10;

QY 1 MKRDLHQFG--PNHGTSGAGSTSSPAVFGKDKMMVKKEE-----DDELLGVIGYKV 52  
DB 1 MKRDHQQFQRLSNHGTSSSSIS-----KDKMMVKKEEDGGNMDDELLAVIGYKV 54  
QY 53 RSSEMAEVALKLEQLETMGMNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELPPAATT 112  
DB 55 RSSEMAEVALKLEQLETMMSNVQEDGLSHLATDTVHYNPSELYSWLDNMLSELNPPPLPA 114  
QY 113 GSNALNPEINNNSNF-FTGGDLKAIPGNACVRSNQFAVDSS-----NKLKPS 166  
DB 115 SSGLDLPVLPSPETCGFPASDYDLKIPGNAI----YQFP-AIDSSSSNNQNRKLSKS 169  
QY 167 SPDSMVTSPSP-----AGVIGTIVTIT-----ESTRPLILVDSQNGVRLVHALMACA 215  
DB 170 SPDSMVTSTGTQIGVIGTIVTITTTTAAESTRSLVDSQNGVRLVHALMACA 229  
QY 216 BAVQSSNLTALAVKQIGFLAVSQAGAMRVKATYFAELARRIYRLSPPTQIDHLSLD 275  
DB 230 EAIQQNNLTALAVKQIGFLAVSQAGAMRVKATYFAELARRIYRLSPQONQIDHCLSD 289  
QY 276 TLQMHFETCYLKFHFTANQALILEAFEGKKRVHVHIDFSMNQGLQWPAALALREGG 335  
DB 290 TLQMHFETCYLKFHFTANQALILEAFEGKKRVHVHIDFSMNQGLQWPAALALREGG 349  
QY 336 PPSRLTIGIPPAADNSDHLHEVGCKLAQALAEIHVEFEYRGFVANSIADLDASMLELRP 395  
DB 350 PPTFRLTIGIPPAADNSDHLHEVGCKLAQALAEIHVEFEYRGFVANSIADLDASMLELRP 409

QY 396 SETEAVAVNSVFELHKLGRITGGIEKVGKQIKPVIPTVVEQESNHNGPVFLDRFTES 455  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 469  
410 SDTEAVAVNSVFELHKLGRPGGIEKVLGVVVKQIKPVIPTVVEQESNHNGPVFLDRFTES 469  
QY 456 LHYYSTLTFDSLEGAPSSQDKVMSEVYLGKQICNLVACEGDDRVVERHETLSQWSNRFGSSG 515  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 515  
470 LHYYSTLTFDSLEGVPNSQDKVMSEVYLGKQICNLVACEGDDRVVERHETLSQWSNRFGSSG 529  
QY 516 FAPAHLSNAPKQASTLLALFNGGEGYRVEKNNGCLMLSMWHTRLPITTSWAKLS-AVH 572  
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 572  
530 LAPAHLSNAPKQASMLLSVFNCGGYRVEESNGLMLGWHTRPLITTSWAKLSTAHAH 587  
RESULT 5  
ADO01803  
ID ADO01803 standard; protein; 587 AA.  
XX  
AC ADO01803;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Thalecress transcription factor protein #108.  
XX  
KW Thalecress; transcription factor; plant; transgenic; abiotic stress;  
KW cold tolerance; heat tolerance; drought; osmotic stress;  
KW phosphate limitation; potassium limitation; nitrogen limitation;  
KW hormone sensitivity; disease resistance; sugar sensing; seed germination;  
KW flowering; inflorescence architectural change;  
KW meristem cell differentiation; phyllotaxy; apical dominance;  
KW trichome development; seed development; premature senescence;  
KW delayed senescence; lethality; necrosis; plant size; leaf morphology;  
KW seed morphology; secondary metabolism; light response; shade avoidance.  
OS Arabidopsis thaliana.  
XX  
PN US2004045049-A1.  
XX  
PD 04-MAR-2004.  
XX  
PF 10-APR-2003; 2003US-00412699.  
XX  
PR 13-SEP-1999; 99US-00394519.  
PR 21-JAN-2000; 2000US-00489376.  
PR 17-FEB-2000; 2000US-00506720.  
PR 22-MAR-2000; 2000US-00532591.  
PR 22-MAR-2000; 2000US-00533029.  
PR 22-MAR-2000; 2000US-00533030.  
PR 22-MAR-2000; 2000US-00533392.  
PR 22-MAR-2000; 2000US-00533648.  
PR 06-APR-2000; 2000WO-US0009448.  
PR 16-NOV-2000; 2000US-00713994.  
PR 17-MAR-2001; 2001US-00819142.  
PR 17-APR-2001; 2001US-00837444.  
PR 30-JAN-2002; 2002US-00958131.  
PR 14-JUN-2002; 2002US-00171468.  
PR 09-AUG-2002; 2002US-00225066.  
PR 09-AUG-2002; 2002US-00225067.  
PR 09-AUG-2002; 2002US-00225068.  
PR 17-DEC-2002; 2002US-0434166P.  
PR 25-FEB-2003; 2003US-00374780.  
XX  
(ZHAN/) ZHANG J.  
PA (FROM/) FROMM M E.  
PA (HEAR/) HEARD J E.  
PA (RIEC/) RIECHMANN J L.  
PA (ADAM/) ADAM L J.  
PA (BROU/) BROUN P E.  
PA (PINE/) PINEDA O.  
PA (KEUD/) REUBER T L.  
PA (KEDD/) KEDDIE J S.  
PA (YUGG/) YU G.  
PA (JIAN/) JIANG C.  
PA (SAMA/) SAMAHA R S.

PA (PILG/) PILGRIM M L.  
 PA (CREE/) CREELMAN R A.  
 PA (DUBE/) DUBELL A N.  
 PA (RATC/) RATCLIFFE O.  
 PA (KUMI/) KUMIMOTO R.  
 PA (SHER/) SHERMAN B K.  
 XX  
 PI Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;  
 PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;  
 PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;  
 PI Sherman BK;  
 XX  
 DR WPI; 2004-225755/21.  
 DR N-PSDB; ADO01802.  
 XX  
 PT New transgenic plant, useful in developing phenotypes with altered or  
 PT improved characteristics or traits.  
 PS Claim 1; SEQ ID NO 216; 213pp; English.  
 XX  
 CC The invention relates to a transgenic plant comprises a recombinant  
 CC polynucleotide having a polynucleotide sequence or its complementary  
 CC sequence comprising a sequence encoding a polypeptide, that initiates  
 CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,  
 CC Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588  
 CC -ADO03527 or ADO03530-ADO03559. Also included are using a transgenic  
 CC plant to grow a progeny plant, an expression cassette (comprising a  
 CC constitutive, inducible or tissue-specific promoter and a recombinant  
 CC polynucleotide described above), a host cell comprising the expression  
 CC cassette, producing a modified plant having a modified trait, identifying  
 CC a factor that is modulated by or interacts with a polypeptide encoded by  
 CC the polynucleotide sequence and identifying at least one downstream  
 CC polynucleotide sequence that is subject to a regulatory effect of any of  
 CC the polypeptides encoded by the polynucleotide described above. The  
 CC transgenic plant is useful for producing a plant that has an altered  
 CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance  
 CC to chilling, germination in cold conditions, freezing tolerance, tolerance  
 CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to  
 CC salt, tolerance to phosphate limitation, tolerance to potassium  
 CC limitation, decreased sensitivity to nitrogen limitation), altered  
 CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered  
 CC response to ethylene, disease resistance, altered susceptibility to  
 CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to  
 CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered  
 CC susceptibility to Sclerotinia, altered sugar sensing, improved seed  
 CC germination and seedling vigor, early flowering, late flowering, extended  
 CC period of flowering, an inflorescence architectural change, a change in  
 CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell  
 CC differentiation, altered phyllotaxy, altered branching pattern, reduced  
 CC apical dominance, reduced trichome density, ectopic trichome development,  
 CC altered trichome development, altered stem morphology, increased root  
 CC growth, increased root hairs, altered seed development, altered cell  
 CC proliferation/cell differentiation, premature senescence, delayed  
 CC senescence, lethality, increased necrosis, an increase in seedling or  
 CC plant size, decreased plant size, a change in leaf morphology, increased  
 CC altered leaf development, increased leaf size and mass, glossy leaves,  
 CC leaf cell expansion, change in seed morphology, altered seed coloration,  
 CC increased seed size, decreased seed size, altered seed shape, change in  
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid  
 CC content, increased leaf insoluble sugars, decreased leaf insoluble  
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid  
 CC content, an alteration of leaf glucosinolate content, change in seed  
 CC biochemistry, an increase in seed oil content, decrease in seed oil  
 CC content, increase in seed fatty acid content, decrease in seed fatty acid  
 CC content, increase in seed protein content, decrease in seed protein  
 CC content, alteration in seed prenyl lipid content, increase in seed  
 CC sterols, upregulation of genes involved in secondary metabolism, increase  
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in  
 CC light response or shade avoidance. The present sequence represents a  
 CC thalacress transcription factor of the invention.  
 XX  
 SQ Sequence 587 AA;

Query Match 81.6%; Score 2408.5; DB 8; Length 587;  
 Best Local Similarity 81.9%; Pred. No. 1.7e-210;  
 Matches 490; Conservative 23; Mismatches 48; Indels 37; Gaps 10;  
 QY 1 MKRDLHOFQG--PNHGTISIAGSTSSPAVFGKDKMMVKKEE-----DDELLGVLGKVK 52  
 DB 1 MKRDHOFQGRLSNHTGSSSSSSIS-----KDKMMVKKEBEDGGNNMDELLAVLGKVK 54  
 QY 53 RSSEMAEVALKLEQLETTMGNAGQEDGLAHLATDTVHYNPAELYSWLDNMLTFLNPPAATT 112  
 DB 55 RSSEMAEVALKLEQLETTMWSNVQEDGLSHLATDTVHYNPSELSYSLWLDNMLSELNPPPLA 114  
 QY 113 GSNALNPEINNNNSP-FTGGDLKATPGNACVRSNQFAFVDS--NKLKPS 166  
 DB 115 SSNGLDPLVPEICGPPASDYDLKVPNGAI-----YQFP-AIDSSSSNNQNKLKCS 169  
 QY 167 SPDGMVTPSP-----AGVIGTIVT-----ESTRLILVDSQDQVRLVHALMACA 215  
 DB 170 SPDGMVTSSTGTGIGVIGTIVT-----TAAAESTRSVILVDSQENGRLVHALMACA 229  
 QY 216 EAVOSSNLTAEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHLSLD 275  
 DB 230 EAIQONNLTLAEALVKQIGCLAVSQAGAMRKVATYFABALARRIYRLSPPOQIDHCLSD 289  
 QY 276 TLOHFYETCPYLKFAHTANQAILAEFGKRVHVIDFSMNQGLQWPAALQALRGG 335  
 DB 290 TLOHFYETCPYLKFAHTANQAILAEFGKRVHVIDFSMNQGLQWPAALQALRGG 349  
 QY 336 PPSRLTIGIPPAADNSDHLHEVGCKLAQAEAHVEFEYGFVANSIADLADSLRLP 395  
 DB 350 PPTRLTIGIPPAADNSDHLHEVGCKLAQAEAHVEFEYGFVANSIADLADSLRLP 409  
 QY 396 SETEAVAVNSVFELHKLGRGTGGIEKRVGVVKQIKPVIPTVVEQSNHNGPVFLDRFTES 455  
 DB 410 SDTEAVAVNSVFELHKLGRPGGIEKVLGVVKQIKPVIPTVVEQSNHNGPVFLDRFTES 469  
 QY 456 LHYYSTLFDLSLEGAPSSQDKVMSEVYLGKQICNLVACGPDPRVERHETLSQNSNRFSSG 515  
 DB 470 LHYYSTLFDLSLEGVPNSQDKVMSEVYLGKQICNLVACGPDPRVERHETLSQNGNRFSSG 529  
 QY 516 FAPAHLSNAFQAOSTLLALFNGGEGYRVEKNKGCLMSWHTRLPITTSAAKLS-AVH 572  
 DB 530 LAPAHLSNAFQAOSMLLSVFNSSGQYRVEESNGCLMGWTRPITTSAAKLSAAH 587  
 RESULT 6  
 AAB28574  
 ID AAB28574 standard; protein; 587 AA.  
 AC AAB28574;  
 XX  
 DT 09-FEB-2001 (first entry)  
 DE Arabidopsis SCLa8.  
 XX  
 KW Scarecrow gene; SCR gene; plant; regulatory gene; SCARECROW-like; SCL;  
 KW transgenic plant; cell division; molecular marker; herbicide resistance;  
 KW salt resistance; pathogen resistance; insect resistance.  
 OS Arabidopsis thaliana.  
 XX  
 PN WO2000053723-A2.  
 XX  
 PD 14-SEP-2000.  
 XX  
 PF 07-MAR-2000; 2000WO-US005875.  
 XX  
 PR 10-MAR-1999; 99US-00265585.  
 XX  
 PA (UUNY ) UNIV NEW YORK STATE.  
 XX  
 PI Benfey PN, Di Laurenzio L, Wysocka-Diller J, Malamy JE, Pysh L;  
 PI Helariutta Y, Bruce W, Lim J;

XX WPI; 2000-594315/56.  
DR N-PSDB; AAC65291.  
XX Scarecrow gene useful for producing transgenic plants expressing genes  
PT whose product increases starch, lignin or cellulose biosynthesis and  
PT confers herbicide, pathogen or insect resistance.  
XX Claim 14; Fig 13; 200pp; English.  
PS  
XX The present sequence is given in a specification relating to the  
CC structure and function of a regulatory gene, SCARECROW (SCR). SCARECROW-  
CC like (SCL) genes encoding proteins containing an amino acid sequence  
CC similar to the sequence of MOT1F III (VHID) of Arabidopsis SCR protein  
CC are disclosed. SCR and SCL genes are useful for producing transgenic  
CC plants whose cell division is modified and root and/or stem development  
CC and gravitropism of stem or hypocotyl is altered. Cell division is  
CC increased or decreased in roots resulting in thicker or thinner root  
CC development. The transgenic plants are useful for expressing a gene of  
CC interest encoding a gene product that confers herbicide, salt, pathogen  
CC or insect resistance in root or embryos and genes encoding starch, lignin  
CC or cellulose biosynthesis in shoots. The SCR gene also confers less  
CC susceptibility to lodging in the transgenic plants than a wild-type  
CC plant. SCR gene sequences are also useful as molecular markers for a  
CC quantitative trait e.g. root or gravitropism trait in molecular breeding  
CC of crop plants  
XX  
SQ Sequence 587 AA;  
  
Query Match 77.7%; Score 2291.5; DB 3; Length 587;  
Best Local Similarity 78.6%; Pred. No. 8.2e-200;  
Matches 470; Conservative 24; Mismatches 67; Indels 37; Gaps 10;  
  
QY 1 MKRDLHOFQ--PNHGTSTAGSTSSPAVFGKDKMMVKKEE-----DDELIGVLGVKY 52  
Db 1 MKRDHQQFGRLSNHTSSSSIS-----KDKMMVKKEEDGGNNMDELLAVLGVKY 54  
  
QY 53 RSSEMAEVALKLEQLETWMSNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELPNPPAAT 112  
Db 55 RSSEMAEVALKLEQLETWMSNAQEDGLSHLATDAAHYNPSELYSWLDNMLSELNPPPLPA 114  
  
QY 113 GSNALNPEINNNNSF-FTGGDLKAIFGNVACRSNQAFADVSS-----NKRLLPSS 166  
Db 115 SNGLDPLPSPFPCGFPXSDYDLKVI PNNAI-----YQFP-AIDSSSSNNQNRKLSCS 169  
  
QY 167 SPDSMVTSPSP-----AGVIGTIVT-----ESTRPLILVDSQDNGVRLVHALMACA 215  
Db 170 SPDSMVTSTGTQIGGVIGTIVTTTTTAAESTRSVILVDSQNGVRLVHALMACA 229  
  
QY 216 EAVQSSNLTALAEALVKQIGFLAVSQAGAMRKVATYFAEALARRIYRSPPTQIDHSLSD 275  
Db 230 EAIQONNLTALAEALVKQIGCLAVSQAGAMRKVATYFAEALARRIYRSPQNQIDHCLSD 289  
  
QY 276 TLQMHFYETCPVLEFAHETANOAILAEPEGKRVHVIDFSMNQGLQWLPALMALREGG 335  
Db 290 TLQMHFYETCPVLEFAHETANOAILAEPEGKRVHVIDFSMNQGLQWLPALMALREGG 349  
  
QY 336 PPSFALTIGIPPAADNSHLHEVGCKLAQAEALHVEFYEGFVANSIADLDASMLELRP 395  
Db 350 PPTFELTGIGPPAPDNSDHLHEVGCKLAQAEALHVEFYEGFVANSIADLDASMLELRP 409  
  
QY 396 SETEAVAVNSVPELHKLGRGTGIEKVGKVIKQICNLVACEGPDPRVERHETLSQWSNRFGSSG 455  
Db 410 SDTEAVAVNSVPELHKLGRGTGIEKVGKVIKQICNLVACEGPDPRVERHETLSQWSNRFGSSG 469  
  
QY 456 LHYSTFLFDSLEGAPSSODKVMSEVYLKQICNLVACEGPDPRVERHETLSQWSNRFGSSG 515  
Db 470 LHYSTXFDSELEGAPSSODKVMSEVYLKQICNLVACEGPDPRVERHETLSQWSNRFGSSG 529  
  
QY 516 FAPAHLSGNAPKQASTLLALFNGGEGYRVEKXNGCLMLSMHTREPLITTSAAKLS-AVH 572  
Db 530 LAPAHLSGNAPKQASMLLSVFNSSQGYRVEBSNCLMLGWHTRPLITTSAAKLSAAH 587

RESULT 7  
AAG38575  
ID AAG38575 standard; protein; 533 AA.  
XX  
AC AAG38575;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47610.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 18-MAY-1999; 99US-0134370P.  
PR 19-MAY-1999; 99US-0134768P.  
PR 20-MAY-1999; 99US-0134911P.  
PR 21-MAY-1999; 99US-0135124P.  
PR 24-MAY-1999; 99US-0135353P.  
PR 25-MAY-1999; 99US-0135629P.  
PR 27-MAY-1999; 99US-0136021P.  
PR 28-MAY-1999; 99US-0136392P.  
PR 01-JUN-1999; 99US-0136782P.  
PR 03-JUN-1999; 99US-0137222P.  
PR 04-JUN-1999; 99US-0137528P.  
PR 08-JUN-1999; 99US-0137502P.  
PR 08-JUN-1999; 99US-0137724P.  
PR 10-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0138847P.  
PR 16-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151085P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151086P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151910P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144335P.	PR 12-OCT-1999;	99US-0158369P.
PR 20-JUL-1999;	99US-0144352P.	PR 13-OCT-1999;	99US-0159293P.
PR 20-JUL-1999;	99US-0144632P.	PR 13-OCT-1999;	99US-0159294P.
PR 20-JUL-1999;	99US-0144884P.	PR 13-OCT-1999;	99US-0159295P.
PR 21-JUL-1999;	99US-0144814P.	PR 14-OCT-1999;	99US-0159329P.
PR 21-JUL-1999;	99US-0145086P.	PR 14-OCT-1999;	99US-0159330P.
PR 21-JUL-1999;	99US-0145088P.	PR 14-OCT-1999;	99US-0159331P.
PR 22-JUL-1999;	99US-0145085P.	PR 14-OCT-1999;	99US-0159637P.
PR 22-JUL-1999;	99US-0145087P.	PR 14-OCT-1999;	99US-0159638P.
PR 22-JUL-1999;	99US-0145089P.	PR 18-OCT-1999;	99US-0159584P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160741P.
PR 23-JUL-1999;	99US-0145145P.	PR 21-OCT-1999;	99US-0160767P.
PR 23-JUL-1999;	99US-0145218P.	PR 21-OCT-1999;	99US-0160768P.
PR 23-JUL-1999;	99US-0145224P.	PR 21-OCT-1999;	99US-0160770P.
PR 26-JUL-1999;	99US-0145276P.	PR 21-OCT-1999;	99US-0160814P.
PR 27-JUL-1999;	99US-0145913P.	PR 21-OCT-1999;	99US-0160815P.
PR 27-JUL-1999;	99US-0145918P.	PR 22-OCT-1999;	99US-0160980P.
PR 27-JUL-1999;	99US-0145919P.	PR 22-OCT-1999;	99US-0160981P.
PR 28-JUL-1999;	99US-0145951P.	PR 25-OCT-1999;	99US-0160989P.
PR 02-AUG-1999;	99US-0146386P.	PR 25-OCT-1999;	99US-0161404P.
PR 02-AUG-1999;	99US-0146388P.	PR 25-OCT-1999;	99US-0161405P.
PR 03-AUG-1999;	99US-0147038P.	PR 25-OCT-1999;	99US-0161406P.
PR 04-AUG-1999;	99US-0147204P.	PR 26-OCT-1999;	99US-0161359P.
PR 05-AUG-1999;	99US-0147302P.	PR 26-OCT-1999;	99US-0161360P.
PR 05-AUG-1999;	99US-0147320P.	PR 26-OCT-1999;	99US-0161361P.
PR 06-AUG-1999;	99US-0147260P.	PR 28-OCT-1999;	99US-0161920P.
PR 06-AUG-1999;	99US-0147303P.	PR 28-OCT-1999;	99US-0161922P.
PR 09-AUG-1999;	99US-0147416P.	PR 28-OCT-1999;	99US-0161993P.
PR 09-AUG-1999;	99US-0147493P.	PR 29-OCT-1999;	99US-0162142P.
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-01493175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		

Query Match 71.6%; Score 2111; DB 3; Length 533;

Best Local Similarity 73.3%; Pred. No. 2.2e-183;

Matches 428; Conservative 35; Mismatches 53; Indels 68; Gaps 8;

Qy	1	MKRDHOFQGNHGTSTAGSSTSSPAVFGKDKMMVKEED-----DELLGVLYKVRSE	56
Db	1	MKRDHOFQGNHGTSTAGSSTSSPAVFGKDKMMVKEED-----DELLGVLYKVRSE	42
Qy	57	MAEVALKLEQLETWGMNAQEDGLAHLATDTVHYNPAELYSLDMMLTEINPPAATTGNA	116
Db	43	MADVAKLEQLEVMNSNVQEDDLSQLATETVHYNPAELYTLWDSMLTDLNPPSS-----	96
Qy	117	LNPEINNNNNNSFFTGGLKAIKPGNVCRRNOFAFVDS-----NKRLLKPS	165
Db	97	-NAEY-----DLKAIPGDAIL---NQFAIDSASSSSNOGGGDTYTTNKRLLKCS	140







PI Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JL;  
 XX Yu G, Samaha R;  
 DR WPI; 2001-335977/35.  
 DR N-PSDB; AAD06646.  
 XX  
 PT Nucleic acids encoding plant transcription factor polypeptides, useful  
 PT for altering the sugar sensing characteristics of plants and increasing  
 PT yield, e.g. corn, potato and cotton plants.  
 XX  
 PS Claim 4; Page 74-76; 151pp; English.  
 XX  
 CC The patent relates to polynucleotides encoding 35 plant transcription  
 CC factors which may be used to modify phenotype associated with a plant's  
 CC sugar sensing characteristics and increasing yield when their expression  
 CC level is altered. Sugars are central regulatory molecules that control  
 CC aspects of physiology, metabolism and development. Therefore the cDNAs  
 CC and proteins of the invention are useful for modifying the growth and  
 CC germination rates of plants, photosynthesis, glyoxylate metabolism,  
 CC respiration, starch and sucrose synthesis and degradation, pathogen  
 CC response, wounding response, cell cycle regulation, pigmentation,  
 CC flowering and senescence of plants and for modifying sink-source  
 CC relationships in seeds, tubers, roots, and other storage organs leading  
 CC to an increase in yield. The transcription factor polynucleotides and  
 CC polypeptides may be used to alter the structure and developmental  
 CC characteristics of plants such as soybean, wheat, corn, potato, cotton,  
 CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,  
 CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,  
 CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,  
 CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,  
 CC tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas.  
 CC The present sequence is an Arabidopsis thaliana transcription factor  
 XX  
 SQ Sequence 532 AA;  
 Query Match 71.4%; Score 2106.5; DB 4; Length 532;  
 Best Local Similarity 73.3%; Pred. No. 5.7e-183;  
 Matches 428; Conservative 36; Mismatches 51; Indels 69; Gaps 9;  
 QY 1 MKRDLHQFGPNIGTSIAGSSTSSPAVFGDKMMVKEED----DELLGVLYKVRSE 56  
 DB 1 MKRDHSH---HHQDK-----KTMNNEEDDGNMDELLAVLYKVRSE 41  
 QY 57 MAEVALKLEQLETMGNAEDGLAHLATDTVHYNPAELYSLDNMLTEINPPAATTGSA 116  
 DB 42 MADVAQKLEQLEVMNSVQEDDLQSLATETVHYNPAELYTMDLSMLTDLNPPSS----- 95  
 QY 117 LNPEINNNNNNSFTGGDLKAIPGNVACRRSNQAFADVSSS-----NKRLLKPS 165  
 DB 96 -NAEY-----DLKAIPGDAIL--NQFAIDSASSNQGGGDTVYTKRLKCS 139  
 QY 166 SSPDSMTVSPSPAGVIGITVITVTESTRPLILVDSQDNGVRLVHALMACAEAVSSNLT 225  
 DB 140 N-----GVETT-TATAESTRHVLVDSQENGVRVHALLACAEAVQKENLT 186  
 QY 226 ABALVKQIGFLAVSQAGAMRKVATVFAELARRIYRLSPPTQIDHSISDTLQWHFYETC 285  
 DB 187 AEALVKQIGFLAVSQIGAMRKVATVFAELARRIYRLSPSPIDHSISDTLQWHFYETC 246  
 QY 286 PYLKFAHTANQAILAEAFEGKRVHVIDFSMNQGLQWPAALMQLALREGGPPSPRLTIG 345  
 DB 247 PYLKFAHTANQAILAEAFQKRVHVIDFSMSQGLQWPAALMQLALRPGPPVPRLTIG 306  
 QY 346 PPAADNSDHLHEVGCKLAELAEATHVEYFGFVANSIADLADSMELRSETEAVAVNS 405  
 DB 307 PPAPNPFYLVHEVGCKLAELAEALHVEFYRGFVANTLADLADSMELRSETEAVAVNS 366  
 QY 406 VFELHKLGRGTGGIEKVGKQKIPVITFVVEESNHNHGVFLDRFTESLHYSTLFD 465  
 DB 367 VFELHKLGRGCAIDKVLGVQVQKPEITFVVEESNHNHSPFLDRFTESLHYSTLFD 426  
 QY 466 LEGAPSSDKQKSEVYLKQICNLVACGDPDRVHERHETLSQWNRFGSAGFAAAHIGSNA 532  
 DB 427 LEGVPSGDKQKSEVYLKQICNLVACGDPDRVHERHETLSQWNRFGSAGFAAAHIGSNA 486  
 QY 526 FKQASTLLALFNGGEGYVERKNNCGMLSWHTRPLITTSAWKLS 569  
 DB 487 FKQASMLLALFNGGEGYVERSDGCLMLGWHTRPLIATSAMKLS 530  
 RESULT 11  
 ADDS5688  
 ID ADDS5688 standard; protein; 532 AA.  
 XX  
 AC ADDS5688;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Thalecress environmental stress-related protein #23.  
 XX  
 KW Thalecress; environmental stress; plant; viral infection;  
 KW fungal infection; microbial infection; herbicide resistance; heat; cold;  
 KW heavy metal; low light; drought; osmotic stress; salt concentration;  
 KW transgenic.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FN US2003131386-A1.  
 PD 10-JUL-2003.  
 XX  
 PF 22-OCT-2002; 2002US-00278536.  
 XX  
 PR 23-MAR-1999; 99US-0125814P.  
 PR 22-MAR-2000; 2000US-00532591.  
 XX  
 PA (SAMA/) SAMAHA R.  
 PA (HEAR/) HEARD J.  
 PA (JIAN/) JIANG C.  
 PA (PINE/) PINEDA O.  
 PA (REUB/) REUBER L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (YUGG/) YU G.  
 PA (KEDD/) KEDDIE J.  
 PA (RATC/) RATCLIFFE O.  
 PA (FILG/) FILGRIM M.  
 PA (ADAM/) ADAM L.  
 PA (BROU/) BROUN P.  
 PA (ZHAN/) ZHANG J.  
 XX  
 PI Samaha R, Heard J, Jiang C, Pineda O, Reuber L, Riechmann JL;  
 PI Yu G, Keddie J, Ratcliffe O, Pilgrim M, Adam L, Broun P, Zhang J;  
 XX  
 DR WPI; 2003-829608/77.  
 DR N-PSDB; ADDS5687.  
 XX  
 PT New recombinant polynucleotide for altering the regulation of gene  
 PT expression of plants to modify the plant's traits, particularly the  
 PT plant's environmental stress tolerance.  
 XX  
 PS Claim 1; SEQ ID NO 46; 219pp; English.  
 XX  
 CC The invention relates to a recombinant polynucleotide that alters a  
 CC plant's environmental stress tolerance when compared with the same trait  
 CC of another plant lacking the recombinant polynucleotide. Also included  
 CC are a transgenic plant comprising the novel recombinant polynucleotide  
 CC having a sequence that encodes a polypeptide comprising at least 6  
 CC consecutive amino acids of any of the 55 250-500 residue amino acid  
 CC sequences (S1), given in the specification, altering the environmental  
 CC stress response or tolerance of a plant, or altering a plant's trait and  
 CC altering the expression levels of at least one gene in a plant. The  
 CC recombinant polynucleotide and methods are useful for altering the  
 CC regulation of gene expression of plants to modify the plant's traits, in  
 CC particular with respect to environmental stress responses (e.g. to viral  
 CC infection, fungal infection, microbial infection, herbicide resistance,  
 CC heat, cold, heavy metals, low light, drought, osmotic stress and salt



QY 526 FKOASTLLALFNGGEGYRVEKNNGCMLMSLWHTPLTTSAWKLS 569  
||||| ||||| ||||| ||||| :||| ||||| ||||| |||||  
Db 487 FKOASMLLALFNGGEGYRVEESDGLMLGWHTRPLIATSAAWKLS 530  
||||| ||||| ||||| ||||| :||| ||||| ||||| |||||  
RESULT 13  
ADI43827  
ID ADI43827 standard; protein; 532 AA.  
XX ADI43827;  
XX ADI43827;  
XX 22-APR-2004 (first entry)  
XX  
XX Plant transcription factor related polypeptide #1502.  
XX transgenic; plant; enhanced tolerance to abiotic stress;  
KW phosphate tolerance; hormone sensitivity; disease resistance;  
KW sugar sensing; flowering; flower structure; stem bifurcation;  
KW branching pattern; apical dominance; trichome; stem morphology;  
KW root growth; root hair; seed development; cell proliferation;  
KW cell differentiation; premature senescence; necrosis; plant size;  
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;  
KW plant anthocyanin; light response; shade avoidance; bioinformatic;  
KW transcription factor; ds.  
XX  
OS Unidentified.  
XX  
XX US2004019927-A1.  
XX  
PD 29-JAN-2004.  
XX  
XX 25-FEB-2003; 2003US-00374780.  
PF  
XX  
PR 18-APR-2001; 2001US-00837944.  
XX  
XX (SHER/) SHERMAN B K.  
PA (RIEG/) RIECHMANN J L.  
PA (JIAN/) JIANG C.  
PA (HEAR/) HEARD J E.  
PA (HAAR/) HAAKE V.  
PA (CREE/) CREELMAN R A.  
PA (RATC/) RATCLIFFE O.  
PA (ADAM/) ADAM L J.  
PA (REUB/) REUBER T L.  
PA (KEDD/) KEDDIE J.  
PA (BROU/) BROUN P E.  
PA (PILG/) PILGRIM M L.  
PA (DUBE/) DUBELL A N.  
PA (PINE/) PINEDA O.  
PA (YUGG/) YU G.  
XX  
XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;  
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;  
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;  
XX  
XX WPI: 2004-132245/13.  
DR N-PSDB; ADI43826.  
XX  
XX New transgenic plant comprising a recombinant polynucleotide of any one  
PT of more than 500 nucleotide sequences, useful in bioinformatic search  
PT methods.  
XX  
XX Disclosure; SEQ ID NO 2290; 435pp; English.  
XX  
XX The invention describes a transgenic plant comprising a recombinant  
CC polynucleotide of any one of more than 500 nucleotide sequences fully  
CC defined in the specification or its complement. The method of the  
CC invention can be used to produce a plant having altered traits such as:  
CC enhanced tolerance to abiotic stress; glyphosate tolerance; hormone  
CC sensitivity; disease resistance; sugar sensing; early or late flowering;  
CC altered flower structure, change in stem bifurcations, altered branching  
CC pattern, reduced apical dominance, reduced trichome density; lack of  
CC

trichomes; reduced ectopic trichome development; altered trichome  
development; increase in trichome number; altered stem morphology;  
increased root growth; increased root hairs; altered seed development;  
altered cell proliferation or cell differentiation; rapid development;  
premature senescence; increased necrosis; increase in seedling or plant  
size; decreased plant size; leaf morphology; seed morphology; seed  
biochemistry; increase in root anthocyanins; increase in plant  
anthocyanins, or alteration in light response or shade avoidance. The  
transgenic plant, polynucleotides and polypeptides are useful in  
bioinformatic search methods. This sequence represents a plant  
transcription factor related polynucleotide.  
XX  
XX Sequence 532 AA;  
XX

Query Match 71.4%; Score 2106.5; DB 8; Length 532;  
Best Local Similarity 73.3%; Pred. No. 5.7e-183;  
Matches 428; Conservative 36; Mismatches 51; Indels 69; Gaps 9;  
QY 1 MKRDLHQFGPNHGTSTAGSSTSPAVFGDKMMVKEED-----DELLGVLYKVSSE 56  
Db 1 MKRDHHH-----HHQDK-----KTMWNEEDDNGMDLAVLYGVKVSSE 41  
QY 57 MAVALKLEQLETHMGNAQEDGLAHLATDVHYNPAELYSWLDNMLTELPPAATTGSNA 116  
Db 42 MADVAQKLEQLEVMMSNVQEDDLSQLATETVHNPABLYTWLDSMLTDLNPPSS----- 95  
QY 117 LNPEINNNNNSPFTGGDLKAIPGNVACRRSNOPAFADVSS-----NKRLLKPS 165  
Db 96 -NAEY-----DLKAIPGDAIL--NQPAIDSASSNQGQGGDYTTNKRUKCS 139  
QY 166 SSPDSMVTSPAGVIGTIVTTTSTRPLILVDSQDNGVRLVHALMACAEVQSSNLT 225  
Db 140 N-----GVVETT-TATAESTRHVVLVDSQENGVRVHALLACAEVQKENTLV 186  
QY 226 AEALVKQIGFLAVSOAGMRKVATYFAEALARIYRLSPQTQIDHLSLDTLQHFYETC 285  
Db 187 AEALVKQIGFLAVSQIGAMEQVATYFAEALARIYRLSPSQSPIDHLSLDTLQHFYETC 246  
QY 286 PYLKFAHFTANQAILAEAFEGKRVHVHIDFSMNOQLWPALMQALALREGGPPFRLTGIG 345  
Db 247 PYLKFAHFTANQAILAEAFQKRVHVHIDFSMSQGLQWLPALMQALALREGGPPFRLTGIG 306  
QY 346 PPAADNSDHLHEVGCKLAQAEAIHVEFEYRGVANSADLDASMLELRPSETEAVAVNS 405  
Db 307 PPAPDNFDYLHEVGCKLAHAEAIHVEFEYRGVANTLADLDASMLELRPSETEAVAVNS 366  
QY 406 VFELHKLGLGTGGIEKVFVGVVKQIKPVIFTVVOESNHNHGPVFLDRFTESLHYSTLFD 465  
Db 367 VFELHKLGLRPGALDKVLGVVNOQKPEIFTVVOESNHNHSPIFLDRFTESLHYSTLFD 426  
QY 466 LEGAPSSQDKVMSEVYLGKQICNLVACEGPDPRVERHETLSQWNRFGSGFAPAHIGSNA 525  
Db 427 LEGVPSQDKVMSEVYLGKQICNVVACDGPDRVERHETLSQWNRFGSGAGFAAHIGSNA 486  
QY 526 FKOASTLLALFNGGEGYRVEKNNGCMLMSLWHTPLTTSAWKLS 569  
Db 487 FKOASMLLALFNGGEGYRVEESDGLMLGWHTRPLIATSAAWKLS 530

RESULT 14  
ADO01805  
ID ADO01805 standard; protein; 532 AA.  
XX  
XX ADO01805;  
XX  
XX 01-JUL-2004 (first entry)  
DT  
XX Thalecress transcription factor protein #109.  
DE  
XX Thalecress; transcription factor; plant; transgenic; abiotic stress;  
KW cold tolerance; heat tolerance; drought; osmotic stress;  
KW phosphate limitation; potassium limitation; nitrogen limitation;  
KW hormone sensitivity; disease resistance; sugar sensing; seed germination;

KW flowering; inflorescence architectural change;  
 KW meristem cell differentiation; phyllotaxy; apical dominance;  
 KW trichome development; seed development; premature senescence;  
 KW delayed senescence; lethality; necrosis; plant size; leaf morphology;  
 KW seed morphology; secondary metabolism; light response; shade avoidance.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FN US2004045049-A1.  
 XX  
 PD 04-MAR-2004.  
 XX  
 XX 10-APR-2003; 2003US-00412699.  
 PF  
 XX 13-SEP-1999; 99US-00394519.  
 PR 17-JAN-2000; 2000US-00489376.  
 PR 21-FEB-2000; 2000US-00506720.  
 PR 22-MAR-2000; 2000US-00532591.  
 PR 22-MAR-2000; 2000US-00533029.  
 PR 22-MAR-2000; 2000US-00533030.  
 PR 22-MAR-2000; 2000US-00533392.  
 PR 22-MAR-2000; 2000US-00533648.  
 PR 06-APR-2000; 2000WO-US009448.  
 PR 16-NOV-2000; 2000US-00713994.  
 PR 27-MAR-2001; 2001US-00819142.  
 PR 17-APR-2001; 2001US-00837444.  
 PR 30-JAN-2002; 2002US-00958131.  
 PR 14-JUN-2002; 2002US-00171468.  
 PR 09-AUG-2002; 2002US-00225066.  
 PR 09-AUG-2002; 2002US-00225067.  
 PR 09-AUG-2002; 2002US-00225068.  
 PR 17-DEC-2002; 2002US-0434166P.  
 PR 25-FEB-2003; 2003US-00374780.  
 XX  
 XX (ZHAN/) ZHANG J.  
 PA (FROM/) FROMM M E.  
 PA (HEAR/) HEARD J B.  
 PA (RIEC/) RIECHMANN J L.  
 PA (ADAM/) ADAM L J.  
 PA (BROU/) BROUN P E.  
 PA (PINE/) PINEDA O.  
 PA (REUB/) REUBER T L.  
 PA (KEDD/) KEDDIE J S.  
 PA (YUGG/) YU G.  
 PA (JIAN/) JIANG C.  
 PA (SAMA/) SAMAHA R S.  
 PA (PILG/) PILGRIM M L.  
 PA (CREE/) CREELMAN R A.  
 PA (DUBE/) DUBELL A N.  
 PA (RATC/) RATCLIFFE O.  
 PA (KUMI/) KUMIMOTO R.  
 PA (SHER/) SHERMAN B K.  
 XX  
 XX Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;  
 PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;  
 PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;  
 PI Sherman BK;  
 XX  
 DR WPI: 2004-225755/21.  
 DR N-PSDB; ADO01804.  
 XX  
 PT New transgenic plant, useful in developing phenotypes with altered or  
 PT improved characteristics or traits.  
 XX  
 XX Claim 1; SEQ ID NO 218; 213pp; English.  
 XX  
 CC The invention relates to a transgenic plant comprises a recombinant  
 CC polynucleotide having a polynucleotide sequence or its complementary  
 CC sequence comprising a sequence encoding a polypeptide, that initiates  
 CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,  
 CC Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588  
 CC -ADO03527 or ADO03530-ADO03559. Also included are using a transgenic  
 CC plant to grow a progeny plant, an expression cassette (comprising a

CC constitutive, inducible or tissue-specific promoter and a recombinant  
 CC polynucleotide described above), a host cell comprising the expression  
 CC cassette, producing a modified plant having a modified trait, identifying  
 CC a factor that is modulated by or interacts with a polypeptide encoded by  
 CC the polynucleotide sequence and identifying at least one downstream  
 CC polynucleotide sequence that is subject to a regulatory effect of any of  
 CC the polypeptides encoded by the polynucleotide described above. The  
 CC transgenic plant is useful for producing a plant that has an altered  
 CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance  
 CC to chilling, germination in cold conditions, freezing tolerance, tolerance  
 CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to  
 CC salt, tolerance to phosphate limitation, tolerance to potassium  
 CC limitation, decreased sensitivity to nitrogen limitation), altered  
 CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered  
 CC response to ethylene, disease resistance, altered susceptibility to  
 CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to  
 CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered  
 CC susceptibility to Sclerotinia, altered sugar sensing, improved seed  
 CC germination and seedling vigor, early flowering, late flowering, extended  
 CC period of flowering, an inflorescence architectural change, a change in  
 CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell  
 CC differentiation, altered phyllotaxy, altered branching pattern, reduced  
 CC apical dominance, reduced trichome density, ectopic trichome development,  
 CC altered trichome development, altered stem morphology, increased root  
 CC growth, increased root hairs, altered seed development, altered cell  
 CC proliferation/cell differentiation, premature senescence, delayed  
 CC senescence, lethality, increased necrosis, an increase in seedling or  
 CC plant size, decreased plant size, a change in leaf morphology, increased  
 CC altered leaf development, increased leaf size and mass, glossy leaves,  
 CC leaf cell expansion, change in seed morphology, altered seed coloration,  
 CC increased seed size, decreased seed size, altered seed shape, change in  
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid  
 CC content, increased leaf insoluble sugars, decreased leaf insoluble  
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid  
 CC content, an alteration of leaf glucosinolate content, change in seed  
 CC biochemistry, an increase in seed oil content, decrease in seed fatty acid  
 CC content, increase in seed fatty acid content, decrease in seed fatty acid  
 CC content, increase in seed protein content, decrease in seed protein  
 CC content, alteration in seed prenyl lipid content, increase in seed  
 CC sterols, upregulation of genes involved in secondary metabolism, increase  
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in  
 CC light response or shade avoidance. The present sequence represents a  
 CC thalacress transcription factor of the invention.  
 XX  
 SQ Sequence 532 AA;

Query Match 71.4%; Score 2106.5; DB 8; Length 532;  
 Best Local Similarity 73.3%; Pred. No. 5.7e-183;

Matches 428; Conservative 36; Mismatches 51; Indels 69; Gaps 9;

Qy 1 MKRDLHQFGQPNHGTSIAGSSTSPAVFGKDKMMVMKEED-----DELLGVLYKVRSE 56  
 Db 1 MKRDHHH-----HHQDK-----KTMWNEEDDGMDELLAVLYKVRSE 41  
 Qy 57 MAEVALKLEQETVMGNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELNPPAATTGSA 116  
 Db 42 MADVAQKLEQLEVMNSVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSS----- 95  
 Qy 117 LNPEINNNNNNSFFTGGDLKAI PGNAVCRNRNQAFADVSSS-----NKRLKPS 165  
 Db 96 -NABY-----DLKAIPGDAIL---NQFAIDSASSNQGGGDTTNNRLKCS 139  
 Qy 166 SSPDSMTVSPAGVIGTGTVTVTSTESTPLILVDSQDNGVRLVHALMACAVQSSNLT 225  
 Db 140 N-----GVVETT-TATAESTFHVVLVDSQENGVRVHALLACAVQKXENLTV 186  
 Qy 226 AEALVKGIFLAVSQAGMRKVATYFAEALARRIYRLSPPTQIDHSLSDTLQMHFYETC 285  
 Db 187 AEALVKGIFLAVSQIGAMRQVATYFAEALARRIYRLSPSQSIDHSLSDTLQMHFYETC 246  
 Qy 286 PYLFAFHTANQALILEAFEGKKRVHVIDFSNQGLOWPALMQALALREGGPPSRLTGIG 345  
 Db 247 PYLFAFHTANQALILEAFQGGKKRVHVIDFSMSQLOWPALMQALALRPGPPVRLTGIG 306



```
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157553P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.

PR 29-OCT-1999; 99US-0162142P.
Query Match 71.3%; Score 2104; DB 3; Length 518;
Best Local Similarity 76.6%; Pred. No. 9.2e-183;
Matches 422; Conservative 34; Mismatches 45; Indels 50; Gaps 7;

QY 34 MMVKEED----DELLGVILGVKVRSSSEMAEVALKLEQLETHWGNQAEQDGLAHLATDTHY 89
DB 1 MMNEEDDNGMDLAVILGVKVRSSSEMAEVALKLEQLEVMMSNVQEDDLSQATETVHY 60
QY 90 NPAELYSWLDNMLTELNPAAATGTSNALNPENNNNNNSFFTGGDLKKAIPGNAVCRRSNQ 149
DB 61 NPAELYTWLDSNMLDLNPPSS-----NAEY-----DLKAIPGDAIL---NQ 98
QY 150 FAFVDDSS-----NKRLKPPSSPDSMTSPSPAGVIGTIVTTVTTESTPLILV 198
DB 99 FAIDSASSNQGGGDYTTTNKRLKCSN-----GVETT-TATAESTRHVILV 145
QY 199 DSQNGVRLVHALMACABAVOSSNLTAEALVKQIGFLAVSQAGAMRKVATYFABALARR 258
DB 146 DSQNGVRLVHALMACABAVQKENLTVAEALVKQIGFLAVSQIGAMRKVATYFABALARR 205
QY 259 IYRLSPQOTQIDHLSLSDTLQMHFYETCPYLKFAHFTANQAILEAPEGKRVHVIDFSMNQ 318
DB 206 IYRLSPQSPIDHLSLSDTLQMHFYETCPYLKFAHFTANQAILEAPEGKRVHVIDFSMSQ 265
QY 319 GLQWALMALALREGPPSPRLTGIGPPAADNSDHLHEVGCKLAQLAEAIHVEFEYRGF 378
DB 266 GLQWALMALALRPGGPPVRLTGIGPPADNFDYLHEVGCKLAHLAEAIHVEFEYRGF 325
QY 379 VANSIADLDASMLELRPSETEAVNSVFLHKLGRGEGIEKVPVGVKQIKPVIPTVVE 438
DB 326 VANTLADLDASMLELRPSEIEAVNSVFLHKLGRPGAIKVLGVVNVQIKPEIFTVVE 385
QY 439 QESNHNGPVFLDRFTESLHYSTLFDLSLEGAPSSODKVMSEVYLGKQICNLVACSGPDRV 498
DB 386 QESNHNSPFLDRFTESLHYSTLFDLSLEGVPSGQDKVMSEVYLGKQICNVVACSGPDRV 445
QY 499 ERHETLSQWNRFGSSGFAPAHLGSNAPKQASTLLALFNGGEGYRVEKNGKCLMLSWHTR 558
DB 446 ERHETLSQWNRFGSAGFAAAHIGSNAPKQASMLLALFNGGEGYRVEESDGLMLGWHTR 505
QY 559 PLITTSAWKLS 569
DB 506 PLIATSARKLS 516

Search completed: November 1, 2004, 21:10:10
Job time : 164 secs
```

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:03:39 ; Search time 42 Seconds  
(without alignments)  
1310.381 Million cell updates/sec

Title: US-10-030-194A-4  
Perfect score: 2950  
Sequence: 1 MKRDLHOFQGNHGTSTAGS.....LSWTRPLTTTSAWKLSAVH 572

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2407	81.6	587	D84426	hypothetical prote
2	2111	71.6	533	H86282	protein F10B6.34 (
3	1535.5	52.1	511	G36688	hypothetical prote
4	1448.5	49.1	523	T51475	RGA-like protein -
5	592	20.1	658	D96656	hypothetical prote
6	588	19.9	653	T47581	SCARECROW1 - Arabi
7	586	19.9	482	E96540	hypothetical prote
8	584	19.8	653	T51244	SCARECROW protein
9	576	19.5	526	E96542	scarecrow-like pro
10	554	18.8	413	G84462	probable SCARECROW
11	516.5	17.5	313	T51242	Scil1 protein (limp
12	507.5	17.2	306	T51236	scarecrow-like pro
13	488	16.5	593	E96347	scarecrow-like 1 p
14	450.5	15.3	352	T51234	scarecrow-like pro
15	438.5	14.9	1336	T02736	probable SCARECROW
16	436	14.8	583	T45597	scarecrow-like pro
17	436	14.8	808	T51232	scarecrow-like pro
18	435	14.7	542	T46142	scarecrow-like 7 (
19	428	14.5	284	T51241	scarecrow-like pro
20	425	14.4	640	B84887	probable SCARECROW
21	423	14.3	375	C71441	probable SCARECROW
22	417	14.1	325	T51235	scarecrow-like pro
23	410.5	13.9	531	T04722	hypothetical prote
24	395.5	13.4	483	T10552	hypothetical prote
25	395.5	13.4	718	T02531	probable SCARECROW
26	390	13.2	573	T51239	scarecrow-like pro
27	358	12.1	623	T47874	scarecrow-like pro
28	350.5	11.9	558	T01343	hypothetical prote
29	344.5	11.7	378	T51237	scarecrow-like pro

30 323.5 11.0 410 2 T45848  
31 307 10.4 486 2 E85433  
32 209 7.1 261 2 T51243  
33 208 7.1 205 2 T51233  
34 177 6.0 133 2 T51240  
35 130 4.4 680 2 T41670  
36 121 4.1 341 2 G59231  
37 117.5 4.0 414 2 S24154  
38 114.5 3.9 1247 2 C89583  
39 114.5 3.9 1742 2 T49451  
40 113.5 3.8 1381 2 S55619  
41 110.5 3.7 414 2 S43253  
42 110.5 3.7 961 2 T23395  
43 110 3.7 836 2 T46070  
44 109.5 3.7 1077 2 T20579  
45 109 3.7 527 2 AI3494

ALIGNMENTS

hypothetical protein At2g01570 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: D84426  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617137  
A:Accession: D84426  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-587 <STO>  
A:Cross-references: UNIPROT:Q9SLH3; GB:AE002093; NID:g3785986; PIDN:AAC67333.1; GSPDB:GN  
C:Genetics:  
A:Gene: At2g01570  
A:Map position: 2

Query Match 81.6%; Score 2407; DB 2; Length 587;  
Best Local Similarity 82.2%; Pred. No. 6.7e-170;  
Matches 488; Conservative 23; Mismatches 47; Indels 36; Gaps 9;

Qy 1 MKRDLHOFQ--PNHGTSTAGSSTSSPAVFGKDKMMVKKEE-----DDELLGVLGK 52  
Db 1 MKRDHOFQGRLSNHGTSSTSSIS-----KDKMMVKKEEDGGNMDDELLAVLGK 54  
Qy 53 RSSEMAEVALKLEQLETMGMNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELPPAAAT 112  
Db 55 RSSEMAEVALKLEQLETMWNVQEDGLSHLATDTVHYNPSELYSWLDNMLSELNPPPLA 114  
Qy 113 GSNALNPEINNNNSP-FTCGDLKALPGNAVCRRSNQFAFVDS--NRKLKSS 166  
Db 115 SSNGLDPLVPSPEICGFPASDYDLKVPGNAI-----YQFP-AIDSSSSNNQKRLKCS 169  
Qy 167 SPDSMTVSPSP-AGVIGTTT---ESTRPLILVDSQDNGVRLVHALMACA 215  
Db 170 SPDSMTVSTGTGIGVGITVTTTTTAAGESTKSVILVDSQENGVRVHALMACA 229  
Qy 216 EAVQSSNLTAEALVKQIGFLAVSQAGAMRVKATVFAEALARRIYRLSPPTQIDHLSLD 275  
Db 230 EAIQONNLTAEALVKQIGCLAVSQAGAMRVKATVFAEALARRIYRLSPQNQIDHCLSD 289  
Qy 276 TLMHFVETCPYLKFAHFTANQAILEAFEGKKRVHVDFSMNQGLQWPAALALREGG 335  
Db 290 TLMHFVETCPYLKFAHFTANQAILEAFEGKKRVHVDFSMNQGLQWPAALALREGG 349  
Qy 336 PPSRLTIGIPPAADNSDHLHEVGCKLAQAEATHVEYRFGVANSIADIDASMLRUP 395  
Db 350 PPTFLRTIGIPPAADNSDHLHEVGCKLAQAEATHVEYRFGVANSIADIDASMLRUP 409

Qy 396 SETEAVAVNSVFELHKLGRGGTEKVFVVVKQIKPVIFVVEQESNHNGPVFLDRFTES 455  
Db 410 SDTEAVAVNSVFELHKLGRGGTEKVLGVVVKQIKPVIFVVEQESNHNGPVFLDRFTES 469  
Qy 456 LHYYSTLFDLSLEGAPSSQDKVMSVYLGGKQICNLVACEGDPDRVERHETLSQWSNRFGSSG 515  
Db 470 LHYYSTLFDLSLEGVPNSQDKVMSVYLGGKQICNLVACEGDPDRVERHETLSQWGNRFGSSG 529  
Qy 516 PAPAHLGSNAPKQASTLLALFNGGEGYRVEKNNGCMLMSWHTRPLIITTSAAWKL 569  
Db 530 LAPAHLGSNAPKQASMLLSVFNSSQGYRVEESNGCMLMGWHTRPLIITTSAAWKL 593

RESULT 2

H86282 protein F10B6.34 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: H86282

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: H86282  
A:Status: preliminary  
A:Molecule type: DNA  
A:Molecule type: DNA  
A:Residues: 1-533 <STO>  
A:Cross-references: UNIPROT:Q9LQT8; GB:AEO05172; NID:g8778219; PIDN:AAF79228.1; GSPDB:CN A:Gene: F10B6.34  
A:Map position: 1

368 VFEHLKLRGCAIDKVLGVWVNIQKPEIFTVVQESNNHNSPIFLDRFTESLHYSTILFDS 427  
 466 LEGAPSSODKVMSEVYLKGQICNLVACGPDPRVERHETLSOWNSRFGSSGFAPAHILGSNA 525  
 428 LEGVPSSGQDKVMSEVYLKGQICNVVACDGPDRVERHETLSQWRNRFGSAGFAAAHIGSNA 487  
 526 FKQASTLIALFNGGSGYRVEKNKGCLMLSWHTRPLITTSAAWKLS 569  
 488 FKQASMLLALFNGGSGYRVEESDGCMLGWHTRPLIATSAAWKLS 531  
 RESULT 3  
 G96688  
 hypothetical protein T27F4.10 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: G96688  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.D.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: G96688  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-511 <STO>  
 A:Cross-references: UNIPROT:Q9C8Y3; GB:AE005173; NID:g10092507; PIDN:AAG12907.1; GSPDB:GN  
 C:Genetics:  
 A:Gene: T27F4.10  
 A:Map position: 1

Qy 479 EYVLGKQICNLVACEGPDVRVHETLSQNSNRFGSGFAPAHLSGNAPKQASTLLALFNG 538  
Db 418 ELFLGRQTLNLVACEGPDVRVHETLSQNSNRFGSGFAPAHLSGNAPKQASTLLALFNG 477  
Qy 539 GEGYRVEKNNCLMLSHWTRPLITTSNAKLSAV 571  
Db 478 ADGYNVEENEGCLLGLWQTRPLITTSNAKLSAV 510

RESULT 4  
T51475  
RGA-like protein - Arabidopsis thaliana  
N:Alternate names: protein K3M16\_60  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: T51475  
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew  
submitted to the Protein Sequence Database, August 2000  
A:Reference number: Z25394  
A:Accession: T51475  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-523 <SAT>  
A:Cross-references: UNIPROT:Q9LF53; EMBL:AL391150  
A:Experimental source: cultivar Columbia; BAC clone K3M16  
C:Genetics:  
A:Map position: 5  
A:Note: K3M16\_60

Query Match 49.1%; Score 1448.5; DB 2; Length 523;  
Best Local Similarity 56.8%; Pred. No. 4.1e-99;  
Matches 304; Conservative 73; Mismatches 111; Indels 47; Gaps 10;  
Qy 38 EEDDELGLVGLYKVRSEMAEVALKLEQETMMGNAQEDGLAHLATDTVHYNPAELYSW 97  
Db 30 DDNDDEFLAVGLYKVRSDMDADVAQKLEQLEWLSNDIASS-SNAFNDTVHYNFSDLSGW 88  
Qy 98 LDNMLTELNPAAATGGSNALNPEINNNNNNSFFTGGDLKAIPGNACRRSQFAVDSS 157  
Db 89 AQSMLSDLN-----YYPOLDPNRI-----CDLRPTDDCECCSN-----SN 125  
Qy 158 SNKRLKSSSPDMSWTSPAGVIGTTVTTESTRPLILVDSQNGVRLVHALMACAEA 217  
Db 126 SNKRIRLGPWCDS-VTS-----ESTRSVLI--EETGVRLVQALVACAEA 167  
Qy 218 VOSSNLTALAEALVKOIGFLAVSQAGMRKVATYFAEALARRIYRLSPQTOIDHSLDPTL 277  
Db 168 VOLENLSLADALVKRVGLLAAQAGAMGVATYFAEALARRIYRHPSAADIDPSFEIL 227  
Qy 278 QMHFYETCPYLKFAHTANQALAEAFEGKRVHVIFDSMNQLOWPALMQALALREGGPP 337  
Db 228 QMNFYDSCPYLKFAHTANQALAEAVTTSRVVHVIDLGNQGMQWPAALMQALALRPGGPP 287  
Qy 338 SFRUTIGPPAADNSDHLHVGCKLAQALAEIHVEFYRGFVANSADLDASMLELRPSE 397  
Db 288 SFRLTGVGNPS--NREGIQELGWKLAQAQAGVBEFKENGLTTERLSLDEPDMFETR-TE 344  
Qy 398 TEAVAVNSVFLHKLGLGTGIEKVGFGVVKQIPVITVVEQSNHNGPVFLDRFTESLH 457  
Db 345 SETLVNVSFVFLHPVLSQFSGIEKLATVAVKAVGLVTVVSEQEANHNGDVFDRFNEALH 404  
Qy 458 YYSTLFDLSLEGAP--SSODKVMSEVYLKQICNLVACEGPDVRVHETLSQNSNRFGSGG 515  
Db 405 YSSLFDSLEGGVITPSQDRVMSEVYLGRQILNLVATGSDRIERHETLAQWRKMGSG 464  
Qy 516 FAPAHLSGNAPKQASTLLALFNGGEGYRVEKNNCLMLSHWTRPLITTSNAKLSA 570  
Db 465 FDPVNLGSDAFKQASLLALSGGGDGYRVEENDGSLMLAWQTGPLIAASAWKLA 519

RESULT 5  
D96656

hypothetical protein F16M19.21 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: D96656  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.;  
ansen, N.P.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: D96656  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-658 <STO>  
A:Cross-references: UNIPROT:Q9CAN3; GB:AE005173; NID:G6598840; PIDN:AAF18695.1; GSPDB:GN  
C:Genetics:  
A:Gene: F16M19.21  
A:Map position: 1

Query Match 20.1%; Score 592; DB 2; Length 658;  
Best Local Similarity 30.3%; Pred. No. 1.2e-35;  
Matches 176; Conservative 71; Mismatches 193; Indels 140; Gaps 21;  
Qy 107 PPAATTGGSNALNPBINNNNN-----SPTTGGD 134  
Db 94 PPLATTRNGEGFWSNNDNNRGSKLKLAEDBSCLSRKRTKCNENGGWFHFTQD 153  
Qy 135 LK--AIPGNACRRSQFAVDSSNNKRLKPS---SSP-----DSMVT----- 173  
Db 154 SSSPALPFLTC-----SGDDEKVCVFPEVISQPLPNWVDSVITELAGIDKQVE 205  
Qy 174 SPSPAGV---IGTTVTVTTESTRPLI-LVDSQDNGVR----- 206  
Db 206 SSLPAAVKEASGGSTSSASSRSLSHRVPEPTNGSRNPYSHRGATEERTTGNINNNNR 265  
Qy 207 -----LVHALMACAEVOSNLTALAEALVKOIGFLAVSQAG--AMRKVATYFAEALA 256  
Db 266 NDLDQDFELVNLGLCLDAIRSNIAINHFARTGDLA-SPRGTPMTRLIAYVIEALA 324  
Qy 257 RRIYRLSP-----PQTQIDHSLD---TLQMHFYETCPYLKFAHTANQALAEAFEGKK 307  
Db 325 LRVARMPHIFHIAAPPREFDRTVEGNSALRFLNQVTPIPKFIHFTANEMLLRAPEGKE 384  
Qy 308 RVHVIDFSMNQLOWPALMQALALREGGPPSFRLTGIGPPAADNSDHLHVGCKLAQALAE 367  
Db 385 RVHVIDFDIKOGLQWPSFFQSLASRINPPHVRITGIG-----ESKLELNETGDRLHGFAB 440  
Qy 368 AIHVEFEYRGFVANSADLDASMLELRPSETEAVAVNSVFLHKL---GRTGGEIKVFGV 425  
Db 441 ANNLQFEPH-VVDRLEDVRLWMLHVK--EGESVAVNCVMQMKTLYDGTGAIRDFGL 497  
Qy 426 VKQIKPVIFTVVEQSNHNGPVFLDRFTESLHYSTLFDLSLEG--APSSQDKV-MSEVYL 482  
Db 498 IRSTNPALVLAQEAQEAHNSQLETRVCNSLKYYSAMFDAIHTNLATDSLARKVKEMLF 557  
Qy 483 GKQICNLVACEGPDVRVHETLSQNSNRFGSGGAPAHLSGNAPKQASTLLALF----- 536  
Db 558 GREIRNIVACSGSRQHRHVRHWRMLEQLGFRSLGSRVERVLQSKMLLRMYGSDNEG 617  
Qy 537 -----NGGGYRVEKNNCLMLSHWTRPLITTSNAW 566  
Db 618 FPNVERSDENGGSGR-----GGGVTLRWSEQPLYTISAW 653

RESULT 6  
T47581  
SCARBCROW1 - Arabidopsis thaliana  
N:Alternate names: protein F24B22.180

C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C;Accession: T47581  
R;Bloeker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: Z23016  
A;Accession: T47581  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-653 <BLO>  
A;Cross-references: UNIPROT:Q9M384; EMBL:AL132957  
A;Experimental source: cultivar Columbia; BAC clone F24B22  
C;Genetics:  
A;Map position: 3  
A;Introns: 511/2  
A;Note: F24B22.180

Query Match 19.9%; Score 588; DB 2; Length 653;  
Best Local Similarity 37.0%; Pred. No. 2.4e-35;  
Matches 158; Conservative 62; Mismatches 165; Indels 42; Gaps 16;  
QY 163 KPSSSDSVMTSPSPAGVIGTIVTTTSTRL-----ILVDSQD-NGVRLVHALMACA 215  
DB 242 RENSSTDA---PPQPTVTATVPAQVNTAEALRERKEIKKQKQDEBGLHLLTLLOCA 298  
QY 216 EAVQSNLTLAEALVKQIGFLAVSQAGAMRKVATFAEALARR-----IVRLSP---- 264  
DB 299 EAVSADNLEANKLLEISQLSTPYGTSAQRAVAAYFSEAMSARLLNSCLIGIYAALPSRW 358  
QY 265 PQTQIDHSL--DTQMHPYETCPYLKFAHTANQAILEAFEGKKRVHVIDFSMQGLQW 322  
DB 359 PQT---HSLKWSAFQV-ENGISPLVKFSHFHTANQAIOEAFEKDSVHIIDIDIMQGLQW 414  
QY 323 PALMQALAREGGPPSFRITGIGPPAADNSDHLHEVGCKLAQAEAIHVEFEYRGFVANS 382  
DB 415 PGLPHILASRPGPGPHVRLTGLG---TSMEAQATGKRLSDFAKGLGPEFCP-LAEK 469  
QY 383 LADLASMELRPSSETEAVVNSVFLHKLGRGTGIEKVGFGVKQIKPVIFTVVEQBSN 442  
DB 470 VGNLDTLRLNR--KREAVAVH--WLQHSLYDVTGSDAHTLWLQLRAPKVVTVVEQDLS 525  
QY 443 HNGPFLDRFTSLHVSYLTFDSL---EGAPSSQDKMSEVYLKQICNLVACGGPDRVE 499  
DB 526 HAGS-FLGRFVBAIHYISALFSLGASYSGESEERHVVVEQQLLSKEIRNLVAVGSPSR-S 583  
QY 500 RHETLSQNSRFGSGGFAPAHGNSAFKQASTLLALFNGGEGYRVEKKNCGICMLSWHTRP 559  
DB 584 GEVKFESMRKMQCGFGKXISLAGNAATQATLLGMF-PSDGYTLVDNDGTILKLGWKDLS 642  
QY 560 LITTSAN 566  
DB 643 LLTASAN 649

RESULT 7  
E96540  
Hypothetical protein Fl1F12.22 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: E96540  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dwyer, K.;  
Hughes, B.; Hulzar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.  
C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: E96540  
A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-482 <STO>  
A;Cross-references: UNIPROT:Q9LPR8; GB:AE005173; NID:G9454566; PIDN:AAF87889.1; GSPDB:GNI  
C;Genetics:  
A;Gene: Fl1F12.22  
A;Map position: 1

Query Match 19.9%; Score 586; DB 2; Length 482;  
Best Local Similarity 31.2%; Pred. No. 2.1e-35;  
Matches 151; Conservative 77; Mismatches 154; Indels 102; Gaps 13;  
QY 170 SMVTSPSPAGVIGTIVTTTSTRLILVDS-----QNGVRLVHALMACAEAV 218  
DB 12 SSVAS-SPLQVFS-----MSLNRPITLASSPPHCLKDLKPEERGLYLHLLTLCANHV 65  
QY 219 OSSNLTAEALVKQIGFLAVSQAGAMRKVATFAEALARRIYR-----LSPPTQID 270  
DB 66 AGSLQANAAALEQLSHLASPGDGTMQRIAAAYFTEALANRLKSWPGLYKALNATQTR-T 124  
QY 271 HSLSDTLQMH--FYETCPYLKFAHTANQAILEAFEGKKRVHVIDFSMQGLQWPAALMQA 328  
DB 125 NVSEERHVRRLFFEMPFIKVSYLLTNRAILLEAMEGEMKVMHVILDLASBPQWLALLOA 184  
QY 329 LALREGPPSFRITGIGPPAADNSDHLHEVGCKLAQAEAIHVEFEYRGFVANSADLDA 388  
DB 185 FNSRPEGPPHLRITGV---HHQKEVLEQMAHRIEAEKLDIPFPQNPVSVR-----LDC 236  
QY 389 SMLB-LRPSETEAVVNSVFLHKLGL-----RTGGIEKVGFGVKQIKPVIFTVVEQBSNHN 446  
DB 237 LNVEQLRVKTEALAVSSVLQLHTFLASDDDLMRKNCALRFQNNPNSGVDLQRLVMMSHGS 296  
QY 415 -----RTGGIEKVGFGVKQIKPVIFTVVEQBSNHN 446  
DB 297 AAARENDNMNNNGYSPSGDSASSLPSPSGRTDSFINAIWGLSPKVMVTEQSDSHGS 356  
QY 447 VFLDRFTSLHYSTLTFDSL---APSSQDKMSEVYL-KQICNLVACGGPDRVERHET 503  
DB 357 TLMERLESYLYVAALFDCLETKVPRTSQDRKVKMLFGEEIKNIISCSEFPERRERHEK 416  
QY 504 LSQNSRFGSGGFAPAHGNSAFKQASTLLALFNGGEGYRVEKKNCGICMLSWHTRPLITT 563  
DB 417 LEKWSQRIDLAGFNGVPLSYAMLQARRLLQ-CGFGDGYRIKEESGCAVICWDRELYSV 475  
QY 564 SAWK 567  
DB 476 SAWR 479

RESULT 8  
T51244  
SCARECROW protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 28-Jul-2000 #sequence\_revision 28-Jul-2000 #text\_change 09-Jul-2004  
C;Accession: T51244  
R;di Laurenzio, L.; Wysocka-Diller, J.; Malamy, J.E.; Pysh, L.; Helariutta, Y.; Freshour,  
submitted to the EMBL Data Library, July 1996  
A;Reference number: Z25338  
A;Accession: T51244  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-653 <DIL>  
A;Cross-references: UNIPROT:Q96304; EMBL:U62798; PIDN:AAB06318.1  
C;Genetics:  
A;Gene: SCARECROW1  
A;Introns: 511/2

Query Match 19.8%; Score 584; DB 2; Length 653;  
Best Local Similarity 36.8%; Pred. No. 4.7e-35;  
Matches 157; Conservative 62; Mismatches 166; Indels 42; Gaps 16;  
QY 163 KPSSSDSVMTSPSPAGVIGTIVTTTSTRL-----ILVDSQD-NGVRLVHALMACA 215  
DB 242 RENSSTDA---PPQPTVTATVPAQVNTAEALRERKEIKKQKQDEBGLHLLTLLOCA 298





Db 368 PESVQSIGGRIITGLRLEQLAEDNGVSFKFKA-MPSKTSIVSPSTLCKPGET--LIVN 424

Qy 405 SVFELHLK-----LGRGTGIEKVGKVKQIKPVFTVVEQESNHNHGPVFLDRFTSLHYYS 460

Db 425 FAFQLHHMPDESVTTVNQRDDELLHMKVSLNPKLTVTVVEQDVNTNTSPFFRFIEAYEYS 484

Qy 461 TLPDSLEGA--PSSQDKV-MSEVYLGKQICNLVACEGPDPRVERHETLSQWNRGSSGFA 517

Db 485 AVFESLDWTLPRESQERNVQRQCLARDIVNIVACEGERIERVEAAGKWRMMAGFN 544

Qy 518 PAHLGSAFQKASTLLALFNGEGYRVEKNNGCLMLSWHTRPLITTSAAK 567

Db 545 KPMGSAKVTNNIQNLKQOYCNK-YLKEEMGELHFCWEKSLIVASAWR 593

RESULT 14

T51234

scarecrow-like protein 1 [imported] - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 28-Jul-2000 #sequence\_revision 28-Jul-2000 #text\_change 09-Jul-2004

C:Accession: T51234

R:Pysh, L.D.; Wysocka-Diller, J.W.; Camilleri, C.; Bouchez, D.; Benfey, P.N.

Plant J. 18, 111-119, 1999

A:Title: The GRAS gene family in Arabidopsis: sequence characterization and basic expression

A:Reference number: Z25337; MUID:99272994; PMID:10341448

A:Accession: T51234

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-352 <PVS>

A:Cross-references: UNIPROT:O9XE51; EMBL:AF036300; PIDN:AAD24403.1

C:Genetics:

A:Gene: SCL1

Query Match 15.3%; Score 450.5; DB 2; Length 352;

Best Local Similarity 32.2%; Pred No. 1.3e-25;

Matches 116; Conservative 70; Mismatches 145; Indels 29; Gaps 12;

Qy 228 ALVQIGFLAVSQAGAMKRVATYFAEALARR-----IYRL-----SPPTQIDHLSLSDT 276

Db 2 SMVNELRQIVSIQDPSQRIAYAMVEGLAARMAASGKFIYALKCKEPPS---DERLA-A 57

Qy 277 LOMHFYETCPYLKPAHFTANQAILAEAFEGKRVHVDFSMNQGLOWPALMQALALREGP 336

Db 58 MQV-LFECVPCPKFGFLAANGAILEAIKEBEVHIIDFINGNQYMTLIRISIAELPGKR 116

Qy 337 PSFRLTGTGPPAA--DNSDHLHEVCCKLAQAEATHVEYRGFVANSIADLADASMLELR 394

Db 117 PRLRUTGIDDESIVORSIGGLRIIGRLLEQLAEDNGVSKFKA-MPSKTSIVSPSTLCK 175

Qy 395 PSETEAVAVNSVFELHKL-----LGRGTGIEKVGKVKQIKPVFTVVEQESNHNHGPVFLD 450

Db 176 PGET--LIVNFAQLHHMPDESVTTVNQRDDELLHMKVSLNPKLTVTVVEQDVNTNTSPFFP 233

Qy 451 RFTSLHYSTLFDLSLEGA--PSSQDKV-MSEVYLGKQICNLVACEGPDPRVERHETLSQW 507

Db 234 RFIEAYEYSVAFESLDMTLPRESQERNVQRQCLARDIVNIVACEGERIERVEAAGK 293

Qy 508 SNRFGSSGFAPAHLSNAFQKASTLLALFNGEGYRVEKNNGCLMLSWHTRPLITTSAAK 567

Db 294 RARMMAGFNPKMPSAKVTNNIQNLKQOYCNK-YLKEEMGELHFCWEKSLIVASAWR 352

RESULT 15

T02736

probable SCARECROW gene regulator [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004

C:Accession: T02736; A84692

R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rong

submitted to the EMBL Data Library, August 1998

A:Description: Arabidopsis thaliana chromosome II BAC T914 genomic sequence.

A:Reference number: Z14710

A:Accession: T02736

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1336 <ROU>

A:Cross-references: UNIPROT:O81074; EMBL:AC005315; NID:g3461834; PID:g3461846

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, P.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: A84692

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1336 <STO>

A:Cross-references: GB:AB002093; NID:g3461846; PIDN:AAC3322.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g29060; T914.14

A:Map position: 2

A:Introns: 694/2

Query Match 14.9%; Score 438.5; DB 2; Length 1336;

Best Local Similarity 28.4%; Pred. No. 8.3e-24;

Matches 132; Conservative 81; Mismatches 177; Indels 75; Gaps 16;

Qy 133 GDLKAIKPGNAVCRSNOPAFAVDSSSNKRLKPSSPDSMTSPSPAG----VIGTTVTTV 188

Db 271 GEAKEQP---VCILNESF-----PKSPAKASTFSKSPKGEKPEASGNSYTK 314

Qy 189 TESTRPLILVDSQDNGVRLVHALMACAEAVOSSNLTAEALVKQIGFLAVSQAGAMRVA 248

Db 315 TPDLRM-----LVSCAQAVINDRRTADELLSRIRQHSSSYGDDGTERLA 359

Qy 249 TYFAEALARRIYRLSPPTQI-----DHSLSDTLMH--FYETCPYLKFAHFTANQAI 299

Db 360 HYFANSEARLAGIG--TQVYTALSSKKTSTSDMLKAYQYIVSVCFFKXIAIIFANHSI 416

Qy 300 --LEAFEGKRVHVDFSMNQGLOWPALMQALALREGGPPSFRLTGTGPPAAD--NSDHL 355

Db 417 MRLASSANAKTIHIIDFGISDGFOWPSLIHLRAWRGSSCKLRITGIELPQGRFPABGV 476

Qy 356 HEVGCKLAQAEATHVEYRGFVANSIADL--DASMLELRPSETEAVAVNSVFELHKL 413

Db 477 IETGRRLAKYCKQFNIPFEY-----NATAQKWSIKLEDLKLKEGEFVAVNSLFRNLL 531

Qy 414 GRTGGI-----EKVFGVVKQIKPVFTVVEQESNHNHGPVFLDRFTSLHYSTLFDLSLEGA 469

Db 532 DETVAVHSPRTVLKIRKIKPDVFIPIGLSGSYNAPFFVTRFREVLPHYSLFDMCDTN 591

Qy 470 PSSQD--KVMSE-VYLGKQICNLVACEGPDPRVERHETLSQWNRGSSGFAPAHLSNAF 526

Db 592 LTREDPMRVMEKEFYGREIMNVVACEGTERVERPESYKQWQARAMRAGFRQIPLEKSLV 651

Qy 527 KQASTLLALFNGGGYR-----VEKNNGCLMLSWHTRPLITTSAAK 566

Db 652 QKLIMVE-----SGYKPKFQVDQDCHWLLQGWKGRIVYGGSIW 691

Search completed: November 1, 2004, 21:14:21

Job time : 44 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:02:40 ; Search time 196 Seconds  
(without alignments)  
1679.155 Million cell updates/sec

Title: US-10-030-194A-4

Perfect score: 2950

Sequence: 1 MKRDHQQFGPNHGTSTAGS.....LSWTRPLITTSWKL SAVH 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2408.5	81.6	587	2 O23642	O23642 arabidopsis
2	2407	81.6	587	2 Q9SLH3	Q9SLH3 arabidopsis
3	2407	81.6	587	2 AAG65090	AAG65090 arabidopsis
4	2402	81.4	587	2 Q941D4	Q941D4 arabidopsis
5	2397	81.3	587	2 O23725	O23725 arabidopsis
6	2111	71.6	533	2 Q9LQ78	Q9LQ78 arabidopsis
7	2110.5	71.5	532	2 O23724	O23724 arabidopsis
8	2106.5	71.4	532	2 O23643	O23643 arabidopsis
9	1906	64.6	590	2 Q8S4W7	Q8S4W7 vitis vinif
10	1869	63.4	587	2 Q6E105	Q6E105 cucurbita m
11	1841.5	62.4	579	2 Q6E106	Q6E106 cucurbita m
12	1761.5	59.7	588	2 Q7T1B6	Q7T1B6 lycopersico
13	1689	57.3	537	2 Q84TQ7	Q84TQ7 gossypium h
14	1602	54.3	630	2 Q9ST48	Q9ST48 zea mays (m
15	1592	54.0	625	2 Q7G7J6	Q7G7J6 oryza sativ
16	1592	54.0	625	2 Q9MB96	Q9MB96 oryza sativ
17	1589.5	53.9	618	2 Q8W127	Q8W127 hordeum vul
18	1568	53.2	623	2 Q9ST59	Q9ST59 triticum ae
19	1565.5	53.1	537	2 Q8S370	Q8S370 argyroxiphi
20	1565	53.1	538	2 Q8S371	Q8S371 argyroxiphi
21	1560	52.9	536	2 Q8S376	Q8S376 dubautia ci
22	1558	52.8	535	2 Q8S369	Q8S369 madia sativ
23	1558	52.8	536	2 Q8S375	Q8S375 dubautia ci
24	1556	52.7	536	2 Q8S374	Q8S374 dubautia me
25	1556	52.7	538	2 Q8RUC4	Q8RUC4 wilkesia gy
26	1553	52.6	539	2 Q8S354	Q8S354 dubautia ra
27	1550.5	52.6	540	2 Q8S365	Q8S365 calycadenia
28	1550	52.5	544	2 Q8S359	Q8S359 argyroxiphi
29	1550	52.5	548	2 Q8S368	Q8S368 cariquistia
30	1549	52.5	547	2 Q9SRP9	Q9SRP9 arabidopsis
31	1545	52.4	538	2 Q8S377	Q8S377 dubautia ar

32 1545 52.4 547 2 O8GXW1 Q8GXW1 arabidopsis  
33 1544 52.3 540 2 Q8S378 Q8S378 dubautia ar  
34 1544 52.3 548 2 Q8S367 Q8S367 cariquistia  
35 1540 52.2 542 2 Q8S360 Q8S360 argyroxiphi  
36 1537.5 52.1 662 2 O65367 O65367 arabidopsis  
37 1536 52.1 537 2 Q8S353 Q8S353 dubautia mi  
38 1536 52.1 537 2 Q8S372 Q8S372 argyroxiphi  
39 1535.5 52.1 511 2 Q9C8Y3 Q9C8Y3 arabidopsis  
40 1535.5 52.1 534 2 Q8S357 Q8S357 dubautia ar  
41 1534 52.0 542 2 Q8S362 Q8S362 argyroxiphi  
42 1531 51.9 538 2 Q8S373 Q8S373 dubautia ra  
43 1530 51.9 537 2 Q8S356 Q8S356 dubautia kn  
44 1529.5 51.8 541 2 Q8S361 Q8S361 argyroxiphi  
45 1528 51.8 537 2 Q8S355 Q8S355 dubautia me

#### ALIGNMENTS

##### RESULT 1

O23642 ID O23642 PRELIMINARY; PRT; 587 AA.  
AC O23642;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE RGAI protein.  
GN Name=RGAI;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97379310; PubMed=9237632;  
RA Truong H.N., Caboche, Daniel-Vedele;  
RT "Sequence and characterization of two Arabidopsis thaliana cDNAs  
isolated by functional complementation of a yeast gln3 gdh1 mutant."  
RL FEBS Lett. 410:213-218(1997).  
DR EMBL; Y11336; CAA72177.1; -.  
DR TRANSFAC; T05385; -.  
DR InterPro; IPR005202; GRAS.  
DR Pfam; PF03514; GRAS; 1.  
SQ SEQUENCE 587 AA; 64023 MW; DOA7A3C741FB51EF CRC64;

Query Match 81.6%; Score 2408.5; DB 2; Length 587;  
Best Local Similarity 81.9%; Pred. No. 8.5e-164;  
Matches 490; Conservative 23; Mismatches 46; Indels 37; Gaps 10;  
Qy 1 MKRDHQQFG--PNHGTSTAGSSTSSPAVFGKDKMMVKKEE-----DDELLGVLYGVK 52  
Db 1 MKRDHQQFGRLSNHGTSSSSSSIS-----KDKMMVKKEEDGGNMDDELLAVLYGVK 54  
Qy 53 RSSEMAEVALKLEQLETMWNAEDGLAHATDTVHYNPALYSWLDNMLTELNPAAAT 112  
Db 55 RSSEMAEVALKLEQLETMWNVQEDGLSHLATDTVHYNPSELYSWLDNMLSELNPPLPA 114  
Qy 113 GSNALNPEINNNNSF--FTGGDLKAIPGNACVRSNOFAFVDS--NKLKPS 166  
Db 115 SSGLDLPVLPSPETCGFPASDYDLKIPGNAI-----YQFP-AIDSSSSNNQKRLKSCS 169  
Qy 167 SPDSNVTSPPSP-----AGVIGTIVTIVT-----ESTRPLILVDSQDQVRLVHALMACA 215  
Db 170 SPDSNVTSSTGTQIGVIGTIVTIVTIVTIVTIVTIVTIVTIVTIVTIVTIVTIVTIVT 229  
Qy 216 EAVGSSNLTALAEVVKQIGFLAVSOAGAMRVATYFAEALARRIYRLSPPTQIDHLSLD 275  
Db 230 EAIQONNLTALAEVVKQIGFLAVSOAGAMRVATYFAEALARRIYRLSPPTQIDHLSLD 289  
Qy 276 TLOMHFYETCPYLKFAHETANQAILAEFGKKVHVDFSMNQGLOWPALMQALREGG 335  
Db 290 TLOMHFYETCPYLKFAHETANQAILAEFGKKVHVDFSMNQGLOWPALMQALREGG 349

```

Qy 336 PPSRLTIGIPPAADNSDHLHEVGCKLAQLAEAHVEFEYRGFVANSIADLADSMLELRP 395
Db 350 PPTFLTGIGPPAPDPSDHLHEVGCKLAQLAEAHVEFEYRGFVANSIADLADSMLELRP 409
Qy 396 SETBAVAVNSVFELHKLGRGTGGIEKVGKQIKPVIPTVVEQESNHNHGVFLDRFTES 455
Db 410 SDTEAVAVNSVFELHKLGRPGGIEKVLGVVQIKPVIPTVVEQESNHNHGVFLDRFTES 469
Qy 456 LHYYSTLFDSELEGAPSSODKVMSEVYLGKQICNLVACGPDPRVERHETLSOWNSRFGSSG 515
Db 470 LHYYSTLFDSELEGAPSSODKVMSEVYLGKQICNLVACGPDPRVERHETLSOWNSRFGSSG 529
Qy 516 FAPAHGSAFQKQASTLLALFNGGEGYRVEKNGKCLMSLWTRPLITTSAMKLS-AVH 572
Db 530 LAPAHGSAFQKQASMLLSVFNSSGGYRVEESNGCLMLGWTRPLITTSAMKLSAAH 587

RESULT 2
ID Q9SLH3 PRELIMINARY; PRT; 587 AA.
AC Q9SLH3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative RGAI, giberellin response modulation protein
DE (At2g01570/F219.19).
GN Name=At2g01570;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
RX [1]
RP SEQUENCE FROM N.A.
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY054160; AAL06821.1; -
DR EMBL; BT010467; AAQ65090.1; -
DR PIR; D84426; D84426.
DR InterPro; IPR005202; GRAS.
DR Pfam; PF03514; GRAS; 1.
SQ SEQUENCE 587 AA; 64035 MW; FC92E7F9408072AA CRC64;

```

Query Match 81.6%; Score 2407; DB 2; Length 587;  
 Best Local Similarity 82.2%; Pred. No. 1.1e-163;  
 Matches 488; Conservative 23; Mismatches 47; Indels 36; Gaps 9;

```

Qy 1 MKRDLHFOG--PNHGTSTIAGSSTSSPAVFGKDKMMVKKEE-----DDELLGVGLGVK 52
Db 1 MKRDLHFOGRLSNHGTSSSSSSIS-----KDKMMVKKEEDGGNMDDELLAVGLGVK 54
Qy 53 RSSMAEVALKLEOLETMMGNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELPPAAAT 112
Db 55 RSSEMAEVALKLEOLETMMGNSVQEDGLSHLATDTVHYNPSELYSWLDNMLSELNPPPLPA 114
Qy 113 GSNALNPEIINNNSNF-FTGGDLKAI PGNAVCHRSNQAFADVDS--NKRLEPSS 166
Db 115 SSGLDLPVLPSPETCGPPASDYDLKVPGNAI-----YQFP-AIDSSSSSSNNQNRKLSKS 169
Qy 167 SPDSMVTSPSP---AGVIGTIVTIVT-----ESTRPLILVDSQDNGVRLVHALMACA 215
Db 170 SPDSMVTSTGTGIGVIGTIVTIVT-----TAAAGESTRSVILVDSQENGVRVHALMACA 229
Qy 216 EAVOSSNLTAELVKQIGFLAVSQAGMRKVATYFAELARRIYRLSPPTQIDHSLSD 275
Db 230 EAIQNNLTAEALVKQIGCLAVSQAGMRKVATYFAELARRIYRLSPQNDHCLSD 289
Qy 276 TLQMHFYETCPYLKFAHFTANQALILEAFEGKRVHVIDFSMNOGLQWPAALQALREGG 335
Db 290 TLQMHFYETCPYLKFAHFTANQALILEAFEGKRVHVIDFSMNOGLQWPAALQALREGG 349
Qy 336 PPSRLTIGIPPAADNSDHLHEVGCKLAQLAEAHVEFEYRGFVANSIADLADSMLELRP 395
Db 350 PPTFLTGIGPPAPDPSDHLHEVGCKLAQLAEAHVEFEYRGFVANSIADLADSMLELRP 409
Qy 396 SETBAVAVNSVFELHKLGRGTGGIEKVGKQIKPVIPTVVEQESNHNHGVFLDRFTES 455
Db 410 SDTEAVAVNSVFELHKLGRPGGIEKVLGVVQIKPVIPTVVEQESNHNHGVFLDRFTES 469
Qy 456 LHYYSTLFDSELEGAPSSODKVMSEVYLGKQICNLVACGPDPRVERHETLSOWNSRFGSSG 515
Db 470 LHYYSTLFDSELEGAPSSODKVMSEVYLGKQICNLVACGPDPRVERHETLSOWNSRFGSSG 529
Qy 516 FAPAHGSAFQKQASTLLALFNGGEGYRVEKNGKCLMSLWTRPLITTSAMKLS 569
Db 530 LAPAHGSAFQKQASMLLSVFNSSGGYRVEESNGCLMLGWTRPLITTSAMKLS 583

RESULT 3
AAQ65090 PRELIMINARY; PRT; 587 AA.
ID AAQ65090
AC AAQ65090;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE At2g01570/F219.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
RX [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT010467; AAQ65090.1; -
DR EMBL; BT010467; AAQ65090.1; -
DR PIR; D84426; D84426.
DR InterPro; IPR005202; GRAS.
DR Pfam; PF03514; GRAS; 1.
SQ SEQUENCE 587 AA; 64035 MW; FC92E7F9408072AA CRC64;

```

Query Match 81.6%; Score 2407; DB 2; Length 587;  
 Best Local Similarity 82.2%; Pred. No. 1.1e-163;  
 Matches 488; Conservative 23; Mismatches 47; Indels 36; Gaps 9;  
 Qy 1 MKRDLHFOG--PNHGTSTIAGSSTSSPAVFGKDKMMVKKEE-----DDELLGVGLGVK 52

```

Db 1 MKRDHQFQGLSHNGTSSSSSSS-----KDKMMVMYKBEDEGGNNMDDDELLAVLGKYV 54
QY 53 RSSEMAEVALKLEOLETWMGNAQEDGLAHLATDTVHYNPAELYSWLDNNMLTELNPAAATT 112
Db 55 RSSEMAEVALKLEOLETWMGNAQEDGLAHLATDTVHYNPSELYSWLDNNMLSELNPPPLPA 114
QY 113 GSNALNPEINNNSNF-FTGGDLKAIPGNVCRRSNOFAFAVDSS-----NKRKLPSS 166
Db 115 SSGLDPLVPSPEICGFPASDYDLKVPGNAI-----YQFP-AIDSSSSNNQNKRLKSCS 169
QY 167 SPDSNVTSPSP-----AGVIGTIVT-----ESTRPLILVDSODNGVRLVHALMACA 215
Db 170 SPDSNVTSTGTQIGGVIPTVTTTTTAAAGESTRSDILVDSQENGVRVHALMACA 229
QY 216 EAVOSSNLTAEALVKQIGFLAVSQAGAMRVATYFAEALARRIYRLSPPTQIDHSLSD 275
Db 230 EAIQONNLTAEALVKQIGFLAVSQAGAMRVATYFAEALARRIYRLSPPTQIDHCLSD 289
QY 276 TLMHFYETCPYLKFAHTANQAILEAFEGKRVHVIDFSMNQGLQWPAALQALALREGG 335
Db 290 TLMHFYETCPYLKFAHTANQAILEAFEGKRVHVIDFSMNQGLQWPAALQALALREGG 349
QY 336 PPSFRLTIGGPPADNSDHLHEVGCKLAQLAEATHVEFEYRGFVANSADLDASMLELRP 395
Db 350 PPTFRLTIGGPPADNSDHLHEVGCKLAQLAEATHVEFEYRGFVANSADLDASMLELRP 409
QY 396 SETEAVAVNSVFELHKLGRGGIEKVGKVKQIKPVIPTVVEQBSNHNPGVFLDRFTES 455
Db 410 SDTEAVAVNSVFELHKLGRGGIEKVGKVKQIKPVIPTVVEQBSNHNPGVFLDRFTES 469
QY 456 LHYYSTLFDLSLEGAPSSQDKVMSEVYLGKQICNLVACGPDPRVERHETLSQNSNRFSGG 515
Db 470 LHYYSTLFDLSLEGAPSSQDKVMSEVYLGKQICNLVACGPDPRVERHETLSQNSNRFSGG 529
QY 516 FAPAHLSNAPKQASTLLALFNGGEGYRVEKNGCMLSMWTRPLITTSANKLS 569
Db 530 LAPAHLSNAPKQASTLLALFNGGEGYRVEKNGCMLSMWTRPLITTSANKLS 583

```

## RESULT 4

```

Q941D4 ID Q941D4 PRELIMINARY; PRT; 587 AA.
AC Q941D4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE At5g01570/F219.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY052239; AAK97709.1; -.
DR InterPro; IPR005202; GRAS.
DR Pfam; PF03514; GRAS; 1.
SQ SEQUENCE 587 AA; 64051 MW; F4490542AAE07670 CRC64;

```

## Query Match

Best Local Similarity 81.4%; Score 2402; DB 2; Length 587;

Matches 488; Conservative 22; Mismatches 48; Indels 36; Gaps 9;

## RESULT 5

```

O23725 ID O23725 PRELIMINARY; PRT; 587 AA.
AC O23725;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GRS protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Columbia;
RA Peng J., Carol P., Richards D.E., King K.E., Cowling R.J.,
RA Murphy G.P., Harberd N.P.;
RT "The Arabidopsis GAI gene defines a signalling pathway that negatively
RT regulates gibberellin responses."
RL Genes Dev. 0:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=Columbia;
RA Harberd N.P.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y15194; CAA75493.1; -.
DR TRANSFAC; T04777; -.
DR InterPro; IPR005202; GRAS.
DR Pfam; PF03514; GRAS; 1.
SQ SEQUENCE 587 AA; 64006 MW; F6F6C7739E7DCA9 CRC64;

```

Query Match		81.3%;	Score 2397;	DB 2;	Length 587;
Best Local Similarity		81.6%;	Pred. No. 5.7e-163;		
Matches		485;	Conservative 25;	Mismatches 48;	Indels 36; Gaps 9;
QY	1 MKRDHLHQFG--PNHGTSIAGSSTSSPAVFGKDKMMVKKEE-----DDELLGVLYGKV	52			
DB	1 MKRDHLHQFGRLSNHGTSSTSSSSIS-----KDKMMVKKEEDGGNNMDELLAVLYGKV	54			
QY	53 RSSEMAEVALKEOLEETMMNGAEOEDGLAHLATDTVHYNPAELYSWLDNMLTELNPPAATT	112			
DB	55 RSSEMAEVALKEOLEETMNSNVQEDGLSHLAADTVHYNPSELYSLWLDNMLSELNPPPLPA	114			
QY	113 GSNALNPEINNNNSF-FTGGDLKATPGNACVRSNOFAFVDS-----NKRLLKPS	166			
DB	115 SSNGLDVLPSPETCGFPASDYDEVIPGNAI----YQFP-AIDSSSSNNQNRLKSCS	169			
QY	167 SPDSWMTSPSP-----AGVIGTIVTTVT-----ESTRPLILVDSQDNGVRLVHALMACA	215			
DB	170 SPDSWMTSTSTGTQIGGVIGTIVTTT-----TAAAESTRSVILVDSQENGVLVHALMACA	229			
QY	216 EAVQSSNLTALBAVKOIGFLAVSOGAMRKVATYFAEALARRIYRLSPPTQIDHSLSD	275			
DB	230 EAIQONNLTALBAVKOIGCLAVSOGAMRKVATYFAEALARRIYRLSPQONQIDHCLSD	289			
QY	276 TLQMHFYETCPYLKFAHTANQAILEAFEGKKRVHVIDFSMNQGLQWPAALMALREGG	335			
DB	290 TLQMHFYETCPYLKFAHTANQAILEAFEGKKRVHVIDFSMNQGLQWPAALMALREGG	349			
QY	336 PPSFRLTGIPPAANDSHLHEVGCKLAQLAEAHVFEYRGFVANSIADLDASMLELRP	395			
DB	350 PPTFRLTGIPPAANDSHLHEVGCKLAQLAEAHVFEYRGFVANSIADLDASMLELRP	409			
QY	396 SETEAVAVNSVPELHKLGRGTGIEKVPVGVKQIKPVIPTVVVEQSNHNGPVFLDRFTES	455			
DB	410 SDTEAVAVNSVPELHKLGRPGIEKVLGVKQIKPVIPTVVVEQSNHNGPVFLDRFTES	469			
QY	456 LHYSTLFDSELAGSPSQDKVMSVYLGKQICNLVACSGPDRVERHETLSQWNRFGSSG	515			
DB	470 LHYSTLFDSELGVPNSQDKVMSVYLGKQICNLVACSGPDRVERHETLSQWNRFGSSG	529			
QY	516 FAPAHLGSNAFKAOSTLLALFNGGVEYRVKNGNGLMLSWHTPLITTSANKLS	569			
DB	530 LAPAHLGSNAFKAQSMLLSVFNQGGYRVESNGLMLGWHTRPLITTSANKUS	583			
RESULT 6					
Q9LQT8					
ID	Q9LQT8	PRELIMINARY;	PRT;	533 AA.	
AC	Q9LQT8;				
DT	01-OCT-2000 (Tremblrel. 15, Created)				
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)				
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)				
DE	F10B6.34 (Atig14920/F10B6.15).				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid2 I; Brassicales; Brassicaceae; Arabidopsis.				
NCBI_TaxID=3702;					
[1]					
RP	SEQUENCE FROM N.A.				
RA	Chao Q., Shinn P., Dunn P., Buehler E., Kahn S., Kim C., Walker M.,				
RA	Williams S., Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A.,				
RA	Hansen N.F., Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S.,				
RA	Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G.,				
RA	Davis R.W., Federpsiel N.A., Theologis A., Ecker J.R.;				
RA	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
[2]					
RP	SEQUENCE FROM N.A.				
RA	Chauk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,				
RA	Khan S., Kim C., Altafi H., Bei B., Chin C., Chieu J., Choi E.,				
RA	Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,				
RA	Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharek N.,				
RA	Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,				

RA	Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,	
RA	Theologis A., Ecker J.;	
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RA	Chouk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,	
RA	Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,	
RA	Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,	
RA	Kawai J., Lam B., Lee J.M., Lin J., Liu S., Miranda M., Narusaka M.,	
RA	Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,	
RA	Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,	
RA	Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,	
RA	Ecker J.R.;	
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.	
RN	[4]	
RP	SEQUENCE FROM N.A.	
RA	Kim C.J., Chen H., Chouk R., Shinn P., Bowser L., Carninci P.,	
RA	Chan M.C., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,	
RA	Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,	
RA	Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,	
RA	Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,	
RA	Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,	
RA	Yamada K., Yu G., Yuan S., Davis R.W., Theologis A., Ecker J.R.;	
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AC006917; AAF79228.1; -	
DR	EMBL; AY058194; AAL25607.1; -	
DR	EMBL; AY142002; AAM98266.1; -	
DR	PIR; H86282; H86282.	
DR	InterPro; IPR005202; GRAS.	
DR	InterPro; IPR009014; Transketo_C_like.	
DR	Pfam; PF03514; GRAS; 1.	
SQ	SEQUENCE 533 AA; 58926 MW; 6464B8C129D79528 CRC64;	
Query Match		
Best Local Similarity 71.6%; Score 2111; DB 2; Length 533;		
Matches 428; Conservative 35; Mismatches 53; Indels 68; Gaps 8;		
QY	1 MKRDHLHQFGPNHGTSIAGSSTSPAVFGKDKMMVKKEE---DELLGVLYGKVRSE 56	
DB	1 MKRDHLHQFGD-----KKTMMNEEDDGNMDELLAVLYGKVRSE 42	
QY	57 MAEVALKEOLETMWNGAEDGLAHLATDTVHYNPAELYSWLDNMLTELNPPAATTGSA 116	
DB	43 MADVAQKLEGLVMSNVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSS----- 96	
QY	117 LNPEINNNNSFFTGDLKAIPGNVACRSNQAFVDS---NKRLLKPS 165	
DB	97 -NAEY-----DLKAIPGDAIL---NQFAIDSASSNQGGDYYTNKRLLKCS 140	
QY	166 SSPDSWMTSPAGVIGTIVTTVTTESTRPLILVDSQDNGVRLVHMACAEVQSSNLT 225	
DB	141 N-----GVVETT-TATAESTRHVVLVDSQENGVRVHALLACAEVQENLT 187	
QY	226 AEAALVQIGFLAVSOGAMRKVATYFAEALARIYRLSPPTQIDHSLSLDTLQMHFYETC 285	
DB	188 AEAALVQIGFLAVSOGAMRKVATYFAEALARIYRLSPSPIDHSLSLDTLQMHFYETC 247	
QY	286 PYLKEAHTANQAILEAFEGKKRVHVIDFSMNQGLQWPAALMALREGGPPSFLRTGIG 345	
DB	248 PYLKEAHTANQAILEAFEGKKRVHVIDFSMSQGLQWPAALMALREGGPPVFRLTGIG 307	
QY	346 PPAADNSDLHEVGCKLAQLAEAHVFEYRGFVANSIADLDASMLELRPSETEAVVNS 405	
DB	308 PPAADNFYDLHEVGCKLAHLAEAHVFEYRGFVANTLADLDASMLELRPSETEAVVNS 367	
QY	406 VFELHKLGRGTGIEKVPVGVKQIKPVIPTVVVEQSNHNGPVFLDRFTESLHYSTLFD 465	
DB	368 VFELHKLGRGTGIEKVPVGVKQIKPVIPTVVVEQSNHNSPFLDRFTESLHYSTLFD 427	
QY	466 LEGAPSSQDKVMSVYLGKQICNLVACSGPDRVERHETLSQWNRFGSSGFAPAHLGSA 525	
DB	428 LEGVPSGDKVMSVYLGKQICNVVACDGPDRVERHETLSQWNRFGSAGFAAAHIGSA 487	



```
QY 466 LEGAPSSQDKWSEVYLGKQICNLVACEGPDVERHETLSQWNRFGSGFAPAHLSGNA 525
D 427 LEGVPSSQDKWSEVYLGKQICNVVACDGPDRVERHETLSQWNRFGSAGFAAAHLSGNA 486
QY 526 FKOASTLLALFNGGEGYRVEKNGCLMLSWHTPLTTSAAKLS 569
D 487 FKOASMLLALFNGGEGYRVEESDGLMGLGWHTPLTTSAAKLS 530

RESULT 9
Q854W7 PRELIMINARY; PRT; 590 AA.
AC Q854W7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE GAI-like protein 1.
GN Name=GAIL;
OS Vitis vinifera (Grape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Vitaceae; Vitis.
OX NCBI_TaxID=29760;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11976683;
RA Boss P.K., Thomas M.R.;
RT "Association of dwarfism and floral induction with a grape 'green
RT revolution' mutation.";
RL Nature 416:847-850(2002).
DR EMBL; AF378125; AM19210.1; -.
DR InterPro; IPR010980; CYC_b562.
DR Pfam; PF03514; GRAS; 1.
SQ SEQUENCE 590 AA; 64866 MW; 0D9CF844C81C0001 CRC64;

Query Match
Best Local Similarity 64.6%; Score 1906; DB 2; Length 590;
Matches 397; Conservative 54; Mismatches 100; Indels 44; Gaps 12;

QY 1 MKRDHLHQFGPNHGTSIAGSSSTSPAVFGKDKMMVKEED---DELLGVLYGKVSSEM 57
D 1 MKREYHH---PHPT-----CSTSEPT--GKGWMDADPQQDAGWDELLAVLYGVNKASDM 50
QY 58 AEVALKLEQLETMGNAQEDGLAHATDTVHYNPAELYSWLDNMLTELNPAAATGSNAL 117
D 51 AEVAQKLEQLEEVIVNAQEDGLSHLASETVHYNPDSLSNWLGLMSLSEFNPTPCALDNPF 110
QY 118 NPEIN-----NNNNNFFETGG-----DLKALPGNAVCERSNQ-----FAFVDSSS 158
D 111 LPPISPLDYNCTQPKQEPSIFDPSLDYDLKALPGKALYSHIEQPPQPPAPPLYORD 170
QY 159 NKRLKSSSPDSMTSPSPAGVIGTPTVTTESTRPLILVDSQNGVRLVHALMACAEV 218
D 171 NKRLKPTTS---ATANSVSSVIGW-GVPTESARPVVLDVDSQGTGLVHTLMACAEV 225
QY 219 QSSNLTALAEVVKQIGFLAVSQAGAMRKVATYFAEALARRIYRLSPQTDHSLSDTLQ 278
D 226 QOENLKLAELVVKQIGFLAVSQAGAMRKVATYFAEALARRIYRLYPDK-PLDSFSFIQ 284
QY 279 MHFYETCPYLKFAHTANQALAEAFEGKRVHVHVDPSNMGLOPMAQALAREGGPPS 336
D 285 MHFYETCPYLKFAHTANQALAEAFEGKRVHVHVDPSNMGLOPMAQALAREGGPPS 344
QY 339 FRLTIGIPPAADNSDHLHEVGCKLAQALAEATHVEFYRGFVANSALDASMLELRPSET 398
D 345 FRLTIGIPPSPTDNTDHLHEVGCKLAQALAEATHVEFYRGFVANSALDASMLELR--DG 402
QY 399 EAVAVNSVFELHKLGRGTGTEKVFVGVVQKIPVFTVVEQSNHNGPVFLDRPTESLHY 458
D 403 ESVAVNSVFELHSLARPGGIERVLSAKOMKPDIVTIVEQANHNGPVFLDRPTESLHY 462
```

```
QY 459 YSTLFDLSLEGA---PSSQDKWSEVYLGKQICNLVACEGPDVERHETLSQWNRFGSS 514
D 463 YSTLFDLSLEGGVGSPVNTQDKLMEVYLGKQICNVVACEGPERVERHETLAQWRARLGA 522
QY 515 GFAPAHLSGNAFKAOSTLLALFNGGEGYRVEKNGCLMLSWHTPLTTSAAKLS 569
D 523 GFDPVNLGSNAFKAQASMLLALFAGGDGYRVEENNGCLMLGWHTPLTTSAAQOLA 577

RESULT 10
Q8E105 PRELIMINARY; PRT; 587 AA.
AC Q8E105;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Gibberellic acid insensitive phloem B.
GN Name=GAIP-B;
OS Cucurbita maxima (Pumpkin) (Winter squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3661;
RN [1]
RP SEQUENCE FROM N.A.
RA Haywood V., Lucas W.J.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY326307; AA096165.1; -.
SQ SEQUENCE 587 AA; 65188 MW; 7BC80F6546F3D81D CRC64;

Query Match
Best Local Similarity 63.4%; Score 1869; DB 2; Length 587;
Matches 383; Conservative 60; Mismatches 98; Indels 64; Gaps 8;

QY 1 MKRDHLHQFGPNHGTSIAGSSSTSPAVFGKDKMMVKEED-----DELLGVLYGKV 52
D 1 MKREHHHLHPRPPDPPSMAAAPNGDTYLTNGKALW---EEDAQLDGGMDLAVLYGKV 56
QY 53 RSSEMAEVALKLEQLETMGNAQEDGLAHATDTVHYNPAELYSWLDNMLTELNPAAAT 112
D 57 KSSDMAEVAQKLEQLEAMCQVDGTGLSHLAFDTVHYNPDSLSTWLESMTLHPHPPSP 116
QY 113 GSNALPPEINNNNNFFETGG-----DLKALPGNAV 143
D 117 -----QPHPSQMDNSFLAPAESSTITSIDYDQRTSSLIFEESSSDYDLKAITSSAI 171
QY 144 CRRSNQFAPAVDSSSNKRLKPSPPDSMTSPSPAGVIGTPTVTTESTRPLILVDSQDN 203
D 172 -----YSPRENKRLKPSSESDDLFSISAIGSNS-----ATRPVLVDSQEN 214
QY 204 GVLVHALMACAEVAVSSNLTALAEALVKQIGFLAVSQAGAMRKVATYFAEALARRIYRLS 263
D 215 GIQLVHALMACAEVAVQNNLNLAELAEKRGYLVAVSQAGAMRKVATFFAEALARRIYRVC 274
QY 264 PQTDHSLSDTLQMHFYETCPYLKFAHTANQALAEAFEGKRVHVHVDPSNMGLOLP 323
D 275 -PENPLDHSMSDMLQHFYESSPYLFAHTANQALAEAFEGKRVHVHVDPSNMGQMP 333
QY 324 ALMQALALREGGPPSPRLTIGIPPAADNSDHLHEVGCKLAQALAEATHVEFYRGFVANS 383
D 334 ALIQLALALPSPGPPAPRLTIGIPPAADNSDYLDQVGWKLAKLVETINVEFYRGFVANS 393
QY 384 ADLDASMLELRPSETEAVAVNSVFELHKLGRGTGTEKVFVGVVQKIPVFTVVEQSNH 443
D 394 ADLDASMLELRPSEVSVVNSVFLHKLARPGAIEKVMVVKQMKPEIMTVVEQANH 453
QY 444 NGPVFLDRPTESLHYSTLFDLSLEGAPSSQDKWSEVYLGKQICNLVACEGPDVERHET 503
D 454 NGPVFLDRPTESLHYSTLFDLSLEGAPSSQDKWSEVYLGKQICNVVACEGSDRVEWHET 513
QY 504 LQSWNRFGSGFAPAHLSGNAFKAOSTLLALFNGGEGYRVEKNGCLMLSWHTPLTTS 563
D 514 LQWTRRLCSSGFEPHILGSLNAFKAQASMLLALFSGGEGYRVEENNGSLTLGWHTPLTIV 573
```

QY 564 SAWKL 568  
AC Q6E106 PRELIMINARY; PRT; 579 AA.  
DB 574 SAWKL 578  
RESULT 11  
Q6E106  
ID AC Q6E106 PRELIMINARY; PRT; 579 AA.  
DT 01-OCT-2004 (TremBLrel. 28, Created)  
DT 01-OCT-2004 (TremBLrel. 28, Last sequence update)  
DT 01-OCT-2004 (TremBLrel. 28, Last annotation update)  
DE Gibberellin acid insensitive phloem.  
GN Name=GAIP;  
OS Cucurbita maxima (Pumpkin) (Winter squash).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.  
OX NCBI\_TaxID=3661;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Haywood V., Lucas W.J.;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY326306; AAQ96164.1; --  
SQ SEQUENCE 579 AA; 64441 MW; 55878C49199C9AD7 CRC64;  
Query Match 62.4%; Score 1841.5; DB 2; Length 579;  
Best Local Similarity 64.2%; Pred. No. 3.6e-123;  
Matches 378; Conservative 64; Mismatches 106; Indels 41; Gaps 8;  
QY 1 MKRDLHQFGPNHGTSIA-GSSTSPAVFGKDKMMVKKEED---DELLGVLGKVRSS 56  
DB 1 MKREHHYLPPEPSPVATGNSRESYLTGKAKLWEEVQLDGGDELLAVLGKVKSSD 60  
QY 57 MAEVALKLEOLETMGNAQEDGLAHLATDVHYNPAELYSMLDNMLTELAPP----- 108  
DB 61 MAEVAQKLEOLEEAMCQVDQGLSHLAFDTVHYNPSDLSTWVESMLTELHPPTSHLDDS 120  
QY 109 -----AATGTSNALNPEINNNN---NNSPFTGGDLKAIPEGNAVCRRSNQFAFAVDSS 158  
DB 121 SFLAPAESSTANVDYEQLOTSSRIPESSSDYDLKAITDSAI-----YSPRE 170  
QY 159 NKRLKSPSSPDMWTSPPAGVIGTVTTVTESTRPLILVDSQNGVRLVHALMACAEAV 218  
DB 171 SKRLKASES-DTDVPSTSAIGASNF-----ATRPVVLVDSQENGIGLVHALMVCAEAV 222  
QY 219 QSSNLTAEALVKQIGFLAVSQAGAMRVATYFAEALARRIYRLSPPTQIDHLSLSDTLQ 278  
DB 223 QONNLTAELVKRIIDYLAVSQAGAMRVATYFAEALARRIYRLC-PENPLDRSVLDMLO 281  
QY 279 MHFYETCPYLKFAHTANQALILEAFEGKRVHVIDFSMNQGLQWPAALALREGGPPS 338  
DB 282 MHFYETCPYLKFAHTANQALILEAFEGKRVHVIDFSMNQGLQWPAALALREGGPT 341  
QY 339 FRLTGIGPAPDNDYLDQVQWKLKFAETLHVEPEYRGFVANSIADLDASMLELRPSET 398  
DB 342 FRLTGIGPAPDNDYLDQVQWKLKFAETLHVEPEYRGFVANSIADLDASMLELRPSEV 401  
QY 399 EAVAVNSVPELHKLGRGTGIEKVGUVKQIKPVIPTVVEQESNHNGPVFLDRFTESLHY 458  
DB 402 ESVVNSVPELHQLAREPGLKVLVSVVKQKPEIPTVVEQEFANNGPVFVERFTESLHY 461  
QY 459 YSTLFDLSLEGAPSSQDKVMSEVILGKQICNLVACGDPDRVERHETLSQWSNRFGSGFAP 518  
DB 462 YSTLFDLSLEGAPSSQDKVMSEVILGKQICNLVACGADPRVERHETLQWRLTSLSAGFDP 521  
QY 519 AHLGSNAFKQASTLLALFNGGEGYRVEKNGCMLSWHTRPLITTSAAWK 567  
DB 522 IHLGSNAFKQASTLLALFNGGEGYRVEKNGCMLSWHTRPLITTSAAWK 570  
RESULT 12

QY1B6  
ID QY1B6 PRELIMINARY; PRT; 588 AA.  
AC QY1B6;  
DT 01-OCT-2003 (TremBLrel. 25, Created)  
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)  
DE GAI-like protein.  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC lamids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bassel G.W., Mullen R.T., Bewley J.D.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY269087; AAP22369.1; --  
DR InterPro; IPR005202; GRAS.  
DR Pfam; PF03514; GRAS; 1  
SQ SEQUENCE 588 AA; 64525 MW; 1EDADAF6BE100621 CRC64;  
Query Match 59.7%; Score 1761.5; DB 2; Length 588;  
Best Local Similarity 60.0%; Pred. No. 2e-117;  
Matches 360; Conservative 72; Mismatches 111; Indels 57; Gaps 8;  
QY 1 MKRDLHQFGPNHGTSIAGSSTSPAVFGKDKMMVKKEED---DELLGVLGKVRSS 55  
DB 1 MKRDRDR-DREKREKRAFSNGAVSS---GSKIKWEEDEERKPDAGMDLAVLGKVKSS 55  
QY 56 EMAEVALKLEOLETMGNAQEDGLAHLATDVHYNPAELYSMLDNMLTELAPP----- 108  
DB 56 DMAVDAQKLEOLEAMGTTMEDGITHLSTDTVHKNPDSMAGWQSVQSMLSSTNFDMCQOE 115  
QY 109 -----AATGTSNALNPEINNNN---NNSPFTGGDLKAIPEGNAVCRRSNQFAFAVDSSSKRLKP 164  
DB 116 NDVLVSGCGSSSIIDFSQNHRTSTISDDDLRAIPGGAV-----FNSDSNKRHS 165  
QY 165 SSSPDSMWTSPPAGVIGTVTTVTESTRPLILVDSQNGVRLVHALMACAEAVQSSNLT 224  
DB 166 TTSSFSFTSS-----SMVTDSASATRPVVLVDSQETGVRLVHTLMACAEAVQENLT 216  
QY 225 LAEALVKQIGFLAVSQAGAMRVATYFAEALARRIYRLSPPTQIDHLSLSDTLQMHFYET 284  
DB 217 LADQLVRHIGILAVSQSGAMRVATYFAEALARRIYKIY-PQDSMESSYTDVLQMHFYET 275  
QY 285 CPYLKFAHTANQALILEAFEGKRVHVIDFSMNQGLQWPAALALREGGPPSPRLTGI 344  
DB 276 CPYLKFAHTANQALILEAFEGKRVHVIDFSMNQGLQWPAALALREGGPPSPRLTGI 335  
QY 345 GPPAADSDHLHVEGCKLAQAEAIHVEFEYRGFVANSIADLDASMLELRPSETSAVAN 404  
DB 336 GPPQPDNTDALQVQWKLKFAETLHVEFEYRGFVANSIADLDATILDRPSETSAVAN 395  
QY 405 SVFELHKLGRGTGIEKVGUVKQIKPVIPTVVEQESNHNGPVFLDRFTESLHYSTLFD 464  
DB 396 SVFELHKLGRGTGIEKVGUVKQIKPVIPTVVEQESNHNGPVFLDRFTESLHYSTLFD 455  
QY 465 SLEGAPSS-----QDKVMSEVILGKQICNLVACGDPDRVERHETLSQWS 508  
DB 456 SLEGAPSS-----QDKVMSEVILGKQICNLVACGDPDRVERHETLSQWS 515  
QY 509 NRGSSGFPAPHLGSNAFKQASTLLALFNGGEGYRVEKNGCMLSWHTRPLITTSAAWK 568  
DB 516 VRMNSSGPDPVHLGSNAFKQASTLLALFNGGEGYRVEKNGCMLSWHTRPLITTSAAWK 575  
RESULT 13  
Q84TQ7  
ID Q84TQ7 PRELIMINARY; PRT; 537 AA.  
AC Q84TQ7;  
DT 01-JUN-2003 (TremBLrel. 24, Created)  
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)





OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,  
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,  
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., VanAken S.E.,  
RA Utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,  
RA White O., Fraser C.M.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Buell R.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC087797; AAKS0137.1; -;  
DR InterPro; IPR005202; GRAS.  
DR Pfam; PF03514; GRAS; 1.  
SQ SEQUENCE 625 AA; 65406 MW; 034FF02719D42E97 CRC64;  
  
Query Match 54.0%; Score 1592; DB 2; Length 625;  
Best Local Similarity 52.5%; Pred. No. 3e-105;  
Matches 341; Conservative 67; Mismatches 137; Indels 104; Gaps 14;  
  
Qy 1 MKRDLHQFGPNHGTSIAGSSTSPAVFGKDKMM--MVKBEED-DELLGVLGKVKRSEM 57  
Db 1 MKREYQEAAGSGSGSSADMGC-----KDKWAGAAGEEEDVELLAALGKVKRSSDM 54  
  
Qy 58 AEVALKLEQLETMG-----NAQEDG-LAHLATDTVHYNPABLYSWLDNMLTELN--- 106  
Db 55 ADVAQKLEQLEWAMGMGVSAFGAADDGFGVSHLATDTVHYNPDLSSWVESMLSELNAPL 114  
  
Qy 107 -----PAAATGSNALNPEINNNNSFFTCGDLKAIPGNACVCRSQFAPAVDSSNK 160  
Db 115 PPIPPAPPAARHAST--SSTVTGGGGGFF-----ELP-----AAADSSSSTY 155  
  
Qy 161 RLKPSPPSPDSMTVTPSPAGVI-----GTTVTVTVTES 191  
Db 156 ALRPISLPVVATADPSAADSARDTKRMETGGSTSSSSSSSLGGGASRGSVVEAAPPA 215  
  
Qy 192 TR-----PLILDSDQNGVRLVHALMACAEAVOSSNLTABALVKOIGFLAVSQA 241  
Db 216 TQGAANAANAPAVPVVVDVTQAGIRLVHALLACAEAVQENFAAAEALVKQIPTILAA SQ 275  
  
Qy 242 GAMRKVATYFAEALARRYRLSP-PQTOIDHSLSDTLQMHFETCPYLKFAHFTANQAIL 300  
Db 276 GAMRKVAAYFGEALARRYRFPADSTLLDAAFADLLHAHFYESCPLYLKFAHFTANQAIL 335  
  
Qy 301 EAFEGKRVHVVIDFSMNQLOWPALMOALALREGGPPSFRLTGIGPPAADNSDHLHEVGC 360  
Db 336 EAFAGCHRVHVVDGFKQGMQWPAALLQALALRPGGPPSFRLTGIGPPDETALQVGVW 395  
  
Qy 361 KLAQLAEAIHVEFEYRGFVANSADLDASMLELR-----PSETEAVAVNSVFELHKLGR 415  
Db 396 KLAQFAHTIRVDFQYRGLVAATLADLEPFMLQPEGEADANEPEVIAVNSVFELHRLIAQ 455  
  
Qy 416 TGGTEKVGUVKQIKYPIVTFVEQESNHPVFLDRFTESLHYSTLFDLSLEGAPSSQ-- 473  
Db 456 PCALEKVLGTVHAVRPRIVTVEQEAHNHNSGFLDRFTESLHYSTMTFDSLEGSSGQAE 515  
  
Qy 474 -----DKVMSVYLKGQICNLVACEGPDPRVERHETLSQNSNRFGSSGFAPAHLL 521  
Db 516 LSPPAAGGGGTDQWMSVYLGRIQICNVVACEGARTERHETLQWRNRLGRAGFEPVHL 575  
  
Qy 522 GSNAPKQASTLLALFNGGEGYRVEKNNGCLMSMHTRELITTSAWKLSA 570  
Db 576 GSNAYKQASTLLALFAGGGRVVEKEGCLTLGWHTRPLIATSAWRVAA 624

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:04:37 ; Search time 41 Seconds  
(without alignments)  
925.217 Million cell updates/sec

Title: US-10-030-194A-4

Perfect score: 2950

Sequence: 1 MKRDLHQFGPNHGTSTAGS.....LSWTRPLITTSANKLSAVH 572

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUTS\_COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2110.5	71.5	532	3	US-09-117-853-2
2	2110.5	71.5	532	4	US-09-911-154-2
3	2110.5	71.5	532	4	US-09-485-529-2
4	2110.5	71.5	532	4	US-09-911-514-2
5	1602	54.3	630	4	US-09-485-529-8
6	1568	53.2	623	4	US-09-485-529-7
7	1416	48.0	630	4	US-09-485-529-1
8	1325	44.9	425	4	US-09-485-529-6
9	1222.5	41.4	277	3	US-09-186-276B-34
10	1222.5	41.4	277	4	US-08-842-445-34
11	1222.5	41.4	277	4	US-09-186-188B-34
12	1089.5	36.9	262	3	US-09-186-276B-35
13	1089.5	36.9	262	4	US-08-842-445-35
14	1089.5	36.9	262	4	US-09-186-188B-35
15	877	29.7	282	3	US-09-117-853-8
16	877	29.7	282	4	US-09-911-154-8
17	877	29.7	282	4	US-09-911-514-8
18	846	28.7	259	3	US-09-117-853-6
19	846	28.7	259	4	US-09-911-154-6
20	846	28.7	259	4	US-09-911-514-6
21	653	22.1	221	3	US-09-117-853-4
22	653	22.1	221	4	US-09-911-154-4
23	653	22.1	221	4	US-09-911-514-4
24	584	19.8	584	3	US-09-186-276B-2
25	584	19.8	584	4	US-08-842-445-2
26	584	19.8	584	4	US-09-186-188B-2
27	516.5	17.5	313	3	US-09-186-276B-27

28	516.5	17.5	313	4	US-08-842-445-27	Sequence 27, Appl
29	516.5	17.5	313	4	US-09-186-188B-27	Sequence 27, Appl
30	507.5	17.2	307	3	US-09-186-276B-31	Sequence 31, Appl
31	507.5	17.2	307	4	US-08-842-445-31	Sequence 31, Appl
32	507.5	17.2	307	4	US-09-186-188B-31	Sequence 31, Appl
33	503.5	17.1	306	3	US-09-186-276B-19	Sequence 19, Appl
34	503.5	17.1	306	4	US-08-842-445-19	Sequence 19, Appl
35	503.5	17.1	306	4	US-09-186-188B-19	Sequence 19, Appl
36	450.5	15.3	353	3	US-09-186-276B-32	Sequence 32, Appl
37	450.5	15.3	353	4	US-08-842-445-32	Sequence 32, Appl
38	450.5	15.3	353	4	US-09-186-188B-32	Sequence 32, Appl
39	439	14.9	351	3	US-09-186-276B-23	Sequence 23, Appl
40	439	14.9	351	4	US-08-842-445-23	Sequence 23, Appl
41	439	14.9	351	4	US-09-186-188B-23	Sequence 23, Appl
42	428	14.5	285	3	US-09-186-276B-54	Sequence 54, Appl
43	428	14.5	285	4	US-08-842-445-54	Sequence 54, Appl
44	428	14.5	285	4	US-09-186-188B-54	Sequence 54, Appl
45	423	14.3	809	3	US-09-186-276B-58	Sequence 58, Appl

#### ALIGNMENTS

RESULT 1

US-09-117-853-2  
; Sequence 2, Application US/09117853  
; Patent No. 6307126  
; GENERAL INFORMATION:  
; APPLICANT: Harberd, Nicholas P  
; APPLICANT: Peng, Jinrong  
; APPLICANT: Carol, Pierre  
; APPLICANT: Richards, Donald E  
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana  
; FILE REFERENCE: 620-45  
; CURRENT APPLICATION NUMBER: US/09/117,853  
; CURRENT FILING DATE: 1998-08-12  
; EARLIER APPLICATION NUMBER: PCT/GB97/00390  
; EARLIER FILING DATE: 1997-02-12  
; EARLIER APPLICATION NUMBER: GB 9602796.6  
; EARLIER FILING DATE: 1996-02-12  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 532  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-117-853-2

Query Match 71.5%; Score 2110.5; DB 3; Length 532;  
Best Local Similarity 73.5%; Pred. No. 6e-192;  
Matches 429; Conservative 35; Mismatches 51; Indels 69; Gaps 9;

Qy	1	MKRD	LHQFGPNHGTSTSPAVFGKDKMMVKEED	----	DELLGVLYKVRSE	56
Db	1	MKRD	HHH-----HHQDK-----	-----	KTMMEEDDNGMDLAVLYKVRSE	41
Qy	57	MAEVA	LKLEQLTWMGNAQEDGLAHLATDTVHYNPAELYSLDNLMTLNPAAVTTGNSA	116		
Db	42	MADVA	KLEQLLEVMSNVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPSS	-----	95	
Qy	117	LNPEI	NNNNNSFTGGDLKAI PGNAVCRNQFAFVDDSS	-----	NKRLKPS	165
Db	96	NAEY	-----DLKAIPGDAIL---	NOFAIDSASSSSNOGGGDTVTTNRLKCS	139	
Qy	166	SSPDS	SVTSPSPAGVIGTIVTTVTRPLIIVDSQDNGVRLVHALMACAEAVQSSNLT	225		
Db	140	N-----	GVVETT-TATAESTRHVVLVDSQENGVRVLRVHALLACAEAVQENLTV	186		
Qy	226	AEALV	KIQIGFLAVSQAGAMRKVATYFAEALARRIYRLSPQQTQIDHLSLDTLQMHFYETC	285		
Db	187	AEALV	KIQIGFLAVSQIGAMRKVATYFAEALARRIYRLSPSQSPIDHLSLDTLQMHFYETC	246		
Qy	286	PYLK	PAHFTANOAILAEAFEGKKRVHIDFSMNQGLQWPAALMQALALREGGPPSFRLTGIG	345		

Db 247 PYLKFHFTANQAILEAFQKKRVHVIDFSMSQGLQWPAALMQALALRPGPPVFLRTGIG 306  
QY 346 PPAADNSDHLHEVGCKLAQAEAIHVEFEYRGFVANSIADLDASMLELRPSETEAVAVNS 405  
Db 307 PPAPDNFDYLHEVGCKLAHAEAIHVEFEYRGFVANTLADLDASMLELRPSEIEISVAVNS 366  
QY 406 VFELHKLGRCAIDKVLGVVQVNIKPFIPTVVEQESNHGPFVFLDRFTESLHYSTLFD 465  
Db 367 VFELHKLGRCAIDKVLGVVQVNIKPFIPTVVEQESNHGPFVFLDRFTESLHYSTLFD 426  
QY 466 LEGAPSSQDKVMSVYLGKQICNLVACGPDPRVERHETLSQWSNRFSGSGPAPAHLSGNA 525  
Db 427 LEGVPSGQDKVMSVYLGKQICNVVACDGPDRVERHETLSQWSNRFSGSAGFAAAHIGSNA 486  
QY 526 FKQASTLLALFNGGEGYRVEKNNGCLMLSWHTRPLITTSAMKLS 569  
Db 487 FKQASMLLALFNGGEGYRVEESDGLMLGWHTRPLIATSAMKLS 530

RESULT 2  
US-09-911-154-2  
; Sequence 2, Application US/09911154  
; Patent No. 6478809  
; GENERAL INFORMATION:  
; APPLICANT: Harberd, Nicholas P  
; APPLICANT: Peng, Jintong  
; APPLICANT: Carol, Pierre  
; APPLICANT: Richards, Donald E  
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana  
; FILE REFERENCE: 620-158  
; CURRENT APPLICATION NUMBER: US/09/911,154  
; CURRENT FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: US 09/117,853  
; PRIOR FILING DATE: 1998-08-12  
; PRIOR APPLICATION NUMBER: PCT/GB97/00390  
; PRIOR FILING DATE: 1997-03-12  
; PRIOR APPLICATION NUMBER: GB 9602796.6  
; PRIOR FILING DATE: 1996-02-12  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 532  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-911-154-2

Query Match 71.5%; Score 2110.5; DB 4; Length 532;  
Best Local Similarity 73.5%; Pred. No. 66-192;  
Matches 429; Conservative 35; Mismatches 51; Indels 69; Gaps 9;  
QY 1 MKRDLHQFGPNHGTSIAGSSTSSPAVFGKDKMMVKEED-----DELLGVLYGKVRSE 56  
Db 1 MKRDHH-----HHQDK-----KTMNNEEDDNGMDLAVLYGKVRSE 41  
QY 57 MAEVALKLEQLETWGMNAQEDGLAHLATDTVHYNPAELYSLWLDNMLTELNPAAATTGNSA 116  
Db 42 MADVAQKLEQLEVMSNVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSS----- 95  
QY 117 LNPEINNNNNNSFFTGDLKAIPGNVACRNSQAFADVSSS-----NKRLKPS 165  
Db 96 -NAEY-----DLKAIPGDAIL---NQFAIDSASSNQGGGDTVTNNKRLKCS 139  
QY 166 SSPDSMTVSPAGVIGTIVTTVTESTRPLILVDSQNGVRLVHALMACAFAVSSNLT 225  
Db 140 N-----GVVETT-TATAESTRHVVLDVSQENGVRVHALLACAAVQKENLTV 186  
QY 226 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC 285  
Db 187 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPSPIDHSLSDTLQMHFYETC 246  
QY 226 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC 285  
Db 187 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPSPIDHSLSDTLQMHFYETC 246  
QY 286 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC 285  
Db 187 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPSPIDHSLSDTLQMHFYETC 246  
QY 286 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC 285

Db 247 PYLKFHFTANQAILEAFQKKRVHVIDFSMSQGLQWPAALMQALALRPGPPVFLRTGIG 306  
QY 346 PPAADNSDHLHEVGCKLAQAEAIHVEFEYRGFVANSIADLDASMLELRPSETEAVAVNS 405  
Db 307 PPAPDNFDYLHEVGCKLAHAEAIHVEFEYRGFVANTLADLDASMLELRPSEIEISVAVNS 366  
QY 406 VFELHKLGRCAIDKVLGVVQVNIKPFIPTVVEQESNHGPFVFLDRFTESLHYSTLFD 465  
Db 367 VFELHKLGRCAIDKVLGVVQVNIKPFIPTVVEQESNHGPFVFLDRFTESLHYSTLFD 426  
QY 466 LEGAPSSQDKVMSVYLGKQICNLVACGPDPRVERHETLSQWSNRFSGSGPAPAHLSGNA 525  
Db 427 LEGVPSGQDKVMSVYLGKQICNVVACDGPDRVERHETLSQWSNRFSGSAGFAAAHIGSNA 486  
QY 526 FKQASTLLALFNGGEGYRVEKNNGCLMLSWHTRPLITTSAMKLS 569  
Db 487 FKQASMLLALFNGGEGYRVEESDGLMLGWHTRPLIATSAMKLS 530

RESULT 3  
US-09-485-529-2  
; Sequence 2, Application US/09485529  
; Patent No. 6762348  
; GENERAL INFORMATION:  
; APPLICANT: Harberd, Nicholas P  
; APPLICANT: Richards, Donald E  
; APPLICANT: Peng, Jintong  
; APPLICANT: Richards, Donald E  
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development  
; FILE REFERENCE: 620-91  
; CURRENT APPLICATION NUMBER: US/09/485,529  
; CURRENT FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB98/02383  
; PRIOR FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: GB 9717192.0  
; PRIOR FILING DATE: 1997-08-13  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 532  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-485-529-2

Query Match 71.5%; Score 2110.5; DB 4; Length 532;  
Best Local Similarity 73.5%; Pred. No. 66-192;  
Matches 429; Conservative 35; Mismatches 51; Indels 69; Gaps 9;  
QY 1 MKRDLHQFGPNHGTSIAGSSTSSPAVFGKDKMMVKEED-----DELLGVLYGKVRSE 56  
Db 1 MKRDHH-----HHQDK-----KTMNNEEDDNGMDLAVLYGKVRSE 41  
QY 57 MAEVALKLEQLETWGMNAQEDGLAHLATDTVHYNPAELYSLWLDNMLTELNPAAATTGNSA 116  
Db 42 MADVAQKLEQLEVMSNVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSS----- 95  
QY 117 LNPEINNNNNNSFFTGDLKAIPGNVACRNSQAFADVSSS-----NKRLKPS 165  
Db 96 -NAEY-----DLKAIPGDAIL---NQFAIDSASSNQGGGDTVTNNKRLKCS 139  
QY 166 SSPDSMTVSPAGVIGTIVTTVTESTRPLILVDSQNGVRLVHALMACAFAVSSNLT 225  
Db 140 N-----GVVETT-TATAESTRHVVLDVSQENGVRVHALLACAAVQKENLTV 186  
QY 226 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC 285  
Db 187 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPSPIDHSLSDTLQMHFYETC 246  
QY 286 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC 285  
Db 187 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPSPIDHSLSDTLQMHFYETC 246  
QY 286 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC 285  
Db 187 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPSPIDHSLSDTLQMHFYETC 246  
QY 286 AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC 285

Db 307 PPAPNFDYLHEVGCKLAHLAIAIHVEFYRGVANTLADLDASMLELRPSIEISVAVNS 366  
Qy 406 VFELHKLGRGTGIEKVGKQIKPVIPTVVEQSNHNGPVFLDRFTESLHYSTLFD 465  
Db 367 VFELHKLGRPCAIDKVLGVNWKIKPEIFTVVEQSNHNSPIFLDRFTESLHYSTLFD 426  
Qy 466 LEGAPSSODKVMSEVYLKQICNLVACGPDORVERHETLSQWNRFGSGGAPAHLSNA 525  
Db 427 LEGVPSGQDKVMSEVYLKQICNVVACDGPDRVERHETLSQWRNFRFGSAGFAAHIGSNA 486  
Qy 526 FQASTLLALFNGGEGYRVEKNNGCLMLSWHTRPLITTSWAKLS 569  
Db 487 FQASMLLALFNGGEGYRVEESDGLMLGWHTRPLIATSAWAKLS 530

## RESULT 4

US-09-911-514-2  
; Sequence 2, Application US/09911514  
; Patent No. 6794560  
; GENERAL INFORMATION:  
; APPLICANT: Harberd, Nicholas P  
; APPLICANT: Peng, Jinrong  
; APPLICANT: Carol, Pierre  
; APPLICANT: Richards, Donald E  
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana  
; FILE REFERENCE: 620-158  
; CURRENT APPLICATION NUMBER: US/09/911,514  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: US 09/117,853  
; PRIOR FILING DATE: 1998-08-12  
; PRIOR APPLICATION NUMBER: PCT/GB97/00390  
; PRIOR FILING DATE: 1997-09-12  
; PRIOR APPLICATION NUMBER: GB 9602796.6  
; PRIOR FILING DATE: 1996-02-12  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 532  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-911-514-2

Query Match 71.5%; Score 2110.5; DB 4; Length 532;  
Best Local Similarity 73.5%; Pred. No. 6e-192;  
Matches 429; Conservative 35; Mismatches 51; Indels 69; Gaps 9;  
Qy 1 MKRDLHQFGPNHGTSTSIAGSSTSSPAVFGKDKMMVKBEED-----DELLGVLYGKVRSE 56  
Db 1 MKRDHH-----HHQDK-----KTMNNEEDDNGMDLALLVLYGKVRSE 41  
Qy 57 MAEVALKLEQLETMGNQAQEDGLAHLATDTVHYNPAELYSMIDNMLTBLNPPAAFTGSNA 116  
Db 42 MADVAQKLEQLEVMNSVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSS----- 95  
Qy 117 LNPEINNNNSFFTGGLKAIPGNAVCRRNQFAFVDDSS-----NKRLLKPS 165  
Db 96 -NABY-----DLKAIPGDAIL---NQFAIDASSNSQGGGDTVTNNKRLKCS 139  
Qy 166 SSPDSMTVSPAGVIGTIVTTVTSTRPLILVDSQDNGVRLVHALMACAEAVQSSNLT 225  
Db 140 N-----GVETT-TATAESTRHVLVDSQENGVRVHALMACAEAVOKENUTV 186  
Qy 226 AEALVKQIGFLAVSQAGAMRKVATYFAEALARRIYRLSPPTQIDHLSLDTLQMHFYETC 285  
Db 187 AEALVKQIGFLAVSQIGAMRKVATYFAEALARRIYRLSPSQSPIDHLSLDTLQMHFYETC 246  
Qy 286 PYLKPFAHTANQAILAEAFEGKKRVHVIDFSMNQGLQWPAALMAQALALRPGGPPVFLTGIG 345  
Db 247 PYLKPFAHTANQAILAEAFQKKRVHVIDFSMSQGLQWPAALMAQALALRPGGPPVFLTGIG 306  
Qy 346 PPAADNSHLHEVGCKLAHLAIAIHVEFYRGVANSIADLDASMLELRPSIETAVANS 405  
Db 307 PPAPNFDYLHEVGCKLAHLAIAIHVEFYRGVANTLADLDASMLELRPSIEISVAVNS 366

Qy 406 VFELHKLGRGTGIEKVGKQIKPVIPTVVEQSNHNGPVFLDRFTESLHYSTLFD 465  
Db 367 VFELHKLGRPCAIDKVLGVNWKIKPEIFTVVEQSNHNSPIFLDRFTESLHYSTLFD 426  
Qy 466 LEGAPSSODKVMSEVYLKQICNLVACGPDORVERHETLSQWNRFGSGGAPAHLSNA 525  
Db 427 LEGVPSGQDKVMSEVYLKQICNVVACDGPDRVERHETLSQWRNFRFGSAGFAAHIGSNA 486  
Qy 526 FQASTLLALFNGGEGYRVEKNNGCLMLSWHTRPLITTSWAKLS 569  
Db 487 FQASMLLALFNGGEGYRVEESDGLMLGWHTRPLIATSAWAKLS 530

## RESULT 5

US-09-485-529-8  
; Sequence 8, Application US/09485529  
; Patent No. 6762348  
; GENERAL INFORMATION:  
; APPLICANT: Harberd, Nicholas P  
; APPLICANT: Richards, Donald E  
; APPLICANT: Peng, Jinrong  
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development  
; FILE REFERENCE: 620-91  
; CURRENT APPLICATION NUMBER: US/09/485,529  
; CURRENT FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB98/02383  
; PRIOR FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: GB 9717192.0  
; PRIOR FILING DATE: 1997-08-13  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 630  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-485-529-8

Query Match 54.3%; Score 1602; DB 4; Length 630;  
Best Local Similarity 52.9%; Pred. No. 2.1e-143;  
Matches 341; Conservative 84; Mismatches 126; Indels 94; Gaps 15;  
Qy 1 MKRDLHQFGPNHGTSTSIAGSSTSSPAVFGKDKMMV-----KERED-DELLGVLYGKVR 53  
Db 1 MKRE-YQDAGSGGD--MGSS-----KDKMAAAGAGQEEDVDLLAALGYKVR 49  
Qy 54 SSEAEVALKLEQLETMWG-----NAQEDGLAHLATDTVHYNPAELYSMIDNMLTE 104  
Db 50 SSDWADVAQKLEQLEMMAMGMGVGAGATADDFVSHLATDTVHYNPSDLSSWVESMLSE 109  
Qy 105 LNPPAA-----TTGSNA-----LNPEINNNNSFFTGGLKAIPG 140  
Db 110 LNAPPALPPATPAPRLASTSTVTSGAAGAGYFDLPVAVDSSSTV-----ALKPIPS 164  
Qy 141 NAVCRSNQFAFVDDSSN-KRLKPSPPDSMTVSPAGVIGTIVTTVTSTR----- 193  
Db 165 PVAAPSADP---STDSAREPKRMRTGGSTSSSSSSSSMDGGRTRSSVVEAAPATQAS 221  
Qy 194 -----PLILVDSQDNGVRLVHALMACAEAVQSSNLTALAEALVKQIGFLAVSQAGAMR 245  
Db 222 AAANGPAPVPPVVDTPQEAIGIRLVHALACAEAVQENFSAEALVKQIPMLASSOGAMR 281  
Qy 246 KVATYFAEALARRIYRLSP--QTQIDHLSLDTLQMHFYETCPYLKFAHTANQAILAEAF 303  
Db 282 KVAAYFGEALARRYVRPPPPDSSILLDAADFALLHAHFYESCPLYKFAHTANQAILAEAF 341  
Qy 304 EGKKRVHVIDFSMNQGLQWPAALMAQALALRPGGPPSFRLTGTGPPAADNSHLHEVGCKLA 363  
Db 342 AGCRRVHVDFGIRKQGMQWPAALQALALRPGGPPSFRLTGTGVPQPDDETALQQYGVKLA 401  
Qy 364 QLAELAIHVEFYRGVANSIADLDASMLELRPSSET-----EAVAVNSVPELHKLGRGTG 419  
Db 402 QFAHTIRVDFQYRGLVAAATLADLBFPMQLQPSGDDTDDDEPVIANSVPELHRLLAQPGAL 461

QY 420 EKVGKVVVKOIKPVITVVEQESNHNCPVFLDRFTESLHYSTLFDLSLEGAPSSQ----- 473  
Db 462 EKVLGTVRAVRPRIVTVVEQEAHNNSGTFLDRFTESLHYSTWFDLSLEGAGSGOSTDA 521  
QY 474 -----DKVMSEVYLGKQICNLVACGPDVRVERHETLSQMSNRFSGSGFAPAHLGNSA 525  
Db 522 SPAAAGGTQVMSEVYLGKQICNVVACGAERTERHETLQWRSLGGSGFAPVHLGNSA 581  
QY 526 FKQASTLLALFNGGEGYRVEKNNGCLMLSWHTRPLITTSAAWLSA 570  
Db 582 YKQASTLLALFAGGDDGYRVEEKXGCLTLGWHTRPLIATSAMRVA 626

RESULT 6  
US-09-485-529-7  
; Sequence 7, Application US/09485529  
; Patent No. 6762348  
; GENERAL INFORMATION:  
; APPLICANT: Harberd, Nicholas P  
; APPLICANT: Richards, Donald E  
; APPLICANT: Peng, Jinrong  
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development  
; FILE REFERENCE: 620-91  
; CURRENT APPLICATION NUMBER: US/09/485,529  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB98/02383  
; PRIOR FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: GB 9717192.0  
; PRIOR FILING DATE: 1997-08-13  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-09-485-529-7

Query Match 53.2%; Score 1568; DB 4; Length 623;  
Best Local Similarity 51.7%; Pred. No. 3.5e-140;  
Matches 336; Conservative 81; Mismatches 123; Indels 110; Gaps 15;

QY 1 MKREDLHQFGPNHGTISAGSTSSPAVFGKDKMMV-----KEEDEDELLGLVGYKVRSS 56  
Db 1 MKREYQDAGSGGGGGGSS-----EDKMMVSAAGGEEVEDELLAALGYKVRASD 52  
QY 57 MAEVALKLEQLFTMW-----GNAQEDGLA-HLATDTVHYNPAELYSMLDNMLTELN-- 106  
Db 53 MADVAQKLEQLMAMGMGVGAGAPDDSFATHLATDVHNPDTLSSWVSMSELNAP 112  
QY 107 -----PPAATGGSNALNPNNNNNNSFTGGDLKAIPGNVCRRSNOPAFADVSSN-KR 161  
Db 113 PPLPLPPA-----PQLNASTSTVTGSGGYFLPP-----SVDSSSSSIYA 151  
QY 162 LKPSSSP-----DSM-----VTSPSPAGVIGTIVTVTES----- 191  
Db 152 LRPIPSAGATAPADLSADSVDRPKRMRTGSGSTSSSSSSSLGGGARSSVVEAAPVPA 211  
QY 192 -----TRPILVDSQNGVRLVHALMACAEAVQSSNLTLAEALVKQIGFLAVSQAGAM 244  
Db 212 AAANATPALPVVVVDTQAGIRLVHALLACAEAVQENLSAAEALVKQIPLLAASQAGAM 271  
QY 245 RKVATYFAEALARRIYRL--SPQQTQIDHSLSDTLQMHFYETCPYLKPAHFTANQAILEA 302  
Db 272 RKVAYFGEALARRVFRFPQDPSLLDAAPADLLHAHFYESCPLYLKAHFTANQAILEA 331  
QY 303 FEGKRVHVIDFSNMQGLQWALMQALREGGPPSFLTGTGPPADNSDHLHEVCCKL 362  
Db 332 FAGCRVHVDFGKQGMQWALQALRPGGPPSFLTGTGVPDPDETALQOVGWKL 391  
QY 363 AQLAEAIHVEFYRGFVANSADLSDASML-----ELRPS-ETEAVAVNSVFLHKLKLGRT 417  
Db 392 AQFAHTIRVDFQYRGLVAATLADLEPFLMQPEGEDPNEEPEVIAVNSVFEMHLLAQPG 451

QY 418 GIEKVGKVVVKOIKPVITVVEQESNHNCPVFLDRFTESLHYSTLFDLSLEG----- 468  
Db 452 ALEKVLGTVRAVRPRIVTVVEQEAHNNSGTFLDRFTESLHYSTWFDLSLEGSGSGGPSE 511  
QY 469 -----APSSODKVMSEVYLGKQICNLVACGPDVRVERHETLSQMSNRFSGSGFAPA 519  
Db 512 VSSGAAAPAAAGTQVMSEVYLGKQICNVVACGAERTERHETLQWRSLGGSGFAPVHLGNSA 571  
QY 520 HLGSAFQKQASTLLALFNGGEGYRVEKNNGCLMLSWHTRPLITTSAAWLS 569  
Db 572 HLGSAFQKQASTLLALFAGGDDGYRVEEKXGCLTLGWHTRPLIATSAMRVA 621

RESULT 7  
US-09-485-529-1  
; Sequence 1, Application US/09485529  
; Patent No. 6762348  
; GENERAL INFORMATION:  
; APPLICANT: Harberd, Nicholas P  
; APPLICANT: Richards, Donald E  
; APPLICANT: Peng, Jinrong  
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development  
; FILE REFERENCE: 620-91  
; CURRENT APPLICATION NUMBER: US/09/485,529  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB98/02383  
; PRIOR FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: GB 9717192.0  
; PRIOR FILING DATE: 1997-08-13  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 630  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
; NAME/KEY: SITE  
; LOCATION: (91)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (94)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (100)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (106)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (118)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (121)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (142)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (197)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (436)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (438)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (474)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (536)  
; OTHER INFORMATION: Xaa is unknown or other amino acid

; NAME/KEY: SITE  
; LOCATION: (558)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (589)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (602)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
; NAME/KEY: SITE  
; LOCATION: (617)  
; OTHER INFORMATION: Xaa is unknown or other amino acid  
US-09-485-529-1

Query Match 48.0%; Score 1416; DB 4; Length 630;  
Best Local Similarity 50.3%; Pred. No. 1.1e-125;  
Matches 320; Conservative 77; Mismatches 155; Indels 84; Gaps 19;

Qy 1 MKRDLHQGFNGHGTISAGSTSSPAVFGKDKMMV-----KEEEDDELLGLVGYKVRSS 56  
Db 10 MKREYQDAGSGGGGGGWS-----EDKMVSAAGGEEVDELLAALGYKVRASD 60  
Qy 57 MAEVALKLEOLETWM-----GNAOEDGLAH-LATDTVHYNPAELYSLDNMLTELN-- 106  
Db 61 MADVAQKLEKLEMAMGCMGVGAGAAPDQVXPXADTVXNPTDXSWVESMLSELXEP 120  
Qy 107 ----PPA-----ATTGNSA---LNPEINNNNNNSFFTGDLKAIPGNVACRRSNQFAFA 153  
Db 121 XPPLPPAPQLNASTVTGGYKDLPPSV--DSSSSIYA---LRPIPSAGATAPAD--LS 173  
Qy 154 VDSSEN-KRLPSSPDPMVTSPAGVIGTIVTTVTES-----TRPLLIVDS 200  
Db 174 ADSVRDPKRMRTGSGSTSS-SSSXSSISLGGGARSSVVEAAPVAAAANATPALPVVVVD 232  
Qy 201 QDNGVRLVHALMACAEAYQSSNLTLAEALVKQIGFLAVSOAGAMRKVATYFAEALARIY 260  
Db 233 QEAGIRLVHALLACAEAYQENLSAAEALVKQIPLLAASOGGAMRKVAAAYFGEALARVF 292  
Qy 261 RL--SPPTQIDHLSLSDTLQMHFETCPYLKFAHTANQALAEAFEGKRVHVIDFSMNQ 318  
Db 293 RFRPQDSSLLDAAPADLLHAHFYESCPLYKFAHTANQALAEAFAGCRRVHVVDGFIQ 352  
Qy 319 GLQWALMAALALREGGPPSFRLTGIGPPADNSDHLHEVCCKLAQLAEATHVFEYRGF 378  
Db 353 GMQWALLQALALRPGGPPSFRLTGIGPPQDPTDALQQVGMKLAQFAHTIRVDVQYRGL 412  
Qy 379 VANSIADLDASML-----ELRPSETEAV-AVNSVRELKLGRTGGIEKVFV-VKQIKPV 432  
Db 413 VAATLADLEPMLQPEGEEDPNEXPVAVNSVFEHRLLAQPGALEKVLGHRAPPCGPE 472  
Qy 433 IFTVVE-QESNHNGPVFLDRFTESLHYSTLFDLSLEG-----APSSQ 473  
Db 473 FXTVVTQEAHNHSGTFLDRFTESLHYSTWFDLSLEGSSGGGPPSEVSSGAAAAPAAAGT 532  
Qy 474 DKWSEVYLKQICNLVACEGPDVRVERHETLSQWNRFGSSGFAPAHIGSNAFKQASTLL 533  
Db 533 DQVXSEVYLGRQICNVVACEGAERTXRETLQWNRNLGNAGFETVHLGNSNAYKQAXTL 592  
Qy 534 ALFNGGEGYRVEKNGCMLSHHTPLITTSANWLS 569  
Db 593 ALFAGGERLXVEEKGCLTGLHTXPLIATSAWRLA 628

RESULT 8  
US-09-485-529-6  
; Sequence 6, Application US/09485529  
; Patent No. 6762348  
; GENERAL INFORMATION:  
; APPLICANT: Harberd, Nicholas P  
; APPLICANT: Richards, Donald E  
; APPLICANT: Peng, Jinrong  
; TITLE OF INVENTION: Genetic Control of Plant Growth and Development  
; FILE REFERENCE: 620-91

; CURRENT APPLICATION NUMBER: US/09/485,529  
; CURRENT FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: PCT/GB98/02383  
; PRIOR FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: GB 9717192.0  
; PRIOR FILING DATE: 1997-08-13  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 425  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-09-485-529-6

Query Match 44.9%; Score 1325; DB 4; Length 425;  
Best Local Similarity 60.5%; Pred. No. 2.5e-117;  
Matches 257; Conservative 54; Mismatches 84; Indels 30; Gaps 5;

Qy 170 SMVTSPAGVIGTIVTTVTETSTPLILVDSODNGVRLVHALMACAEAYQSSNLTLAEAL 229  
Db 4 SVVEAAPVAAAANATPAL-----PVVVVDTEAGIRLVHALLACAEAYQENLSAAEAL 58  
Qy 230 VKQIGFLAVSOAGAMRKVATYFAEALARIYRL--SPPTQIDHLSLSDTLQMHFETCPY 287  
Db 59 VKQIPLLAASOGGAMRKVAAAYFGEALARVFRPQDPSLLDAFADLLHAHFYESCYP 118  
Qy 288 LKFAHTANQALAEAFEGKRVHVIDFSMNQGLQWALMAALREGGPPSFRLTGIGPP 347  
Db 119 LKFAHTANQALAEAFAGCRRVHVVDGFIKQGMQWALLQALALRPGGPPSFRLTGIGPP 178  
Qy 348 AADNSDHLHEVCCKLAQLAEATHVFEYRGVANSIADLDASML-----ELRPS-ETEAVA 402  
Db 179 QPDETDLAQVGMKLAQFAHTIRVDVQYRGLVAAATLADLEPMLQPEGEEDPNESPEVIA 238  
Qy 403 VNSVFEHLKLGRTGGIEKVFVGVKQIKPVLTFTVVEQSNHNGPVFLDRFTESLHYSTL 462  
Db 239 VNSVFEHRLLAQPGALEKVLGTVRAVPRIVTVVEQEAHNHSGTFLDRFTESLHYSTM 298  
Qy 463 FDSLEG-----APSSQDKWSEVYLKQICNLVACEGPDVRVERHETL 504  
Db 299 FDSLEGSSGGGPPSEVSSGAAAAPAAAGTDQWSEVYLGRQICNVVACEGAERTERHETL 358  
Qy 505 SQWNRFGSSGFAPAHIGSNAFKQASTLLALFNGGEGYRVEKNGCMLSHHTPLITTS 564  
Db 359 GQWNRNLGNAGFETVHLGNSNAYKQASTLLALFAGDGYKVEEKGCLTGLMHTRLIATS 418  
Qy 565 ANWLS 569  
Db 419 ANWLA 423

RESULT 9  
US-09-186-276B-34  
; Sequence 34, Application US/09186276B  
; Patent No. 6388173  
; GENERAL INFORMATION:  
; APPLICANT: Benfey, Philip  
; APPLICANT: DiLaurenzio, Laura  
; APPLICANT: Wysocka-Diller, Joanna  
; APPLICANT: Malamy, Jocelyn E.  
; APPLICANT: Fysh, Leonard  
; APPLICANT: Helariutta, Yrjo  
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses Thereof  
; FILE REFERENCE: 5914-075-999  
; CURRENT APPLICATION NUMBER: US/09/186,276B  
; CURRENT FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: 08/842,445  
; PRIOR FILING DATE: 1997-04-24  
; PRIOR APPLICATION NUMBER: 08/638,617  
; PRIOR FILING DATE: 1996-04-26  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 34

```

; LENGTH: 277
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(277)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-186-276B-34

Query Match
Best Local Similarity 41.4%; Score 1222.5; DB 3; Length 277;
Matches 246; Conservative 6; Mismatches 14; Indels 11; Gaps 2;

QY 159 NKRLKSPSSPDSMTSPSP-----AGVIGTIVTTVT-----ESTRPLILVDSQDNGVRL 207
Db 1 NKRLKSCSPDSMTSTGTQIGGVIGTIVTTTITTTTAAAESTRSVILVDSQENGVR 60

QY 208 VHALMACAEAVQSSNLTAEALVKQIGFLAVSQAGMRKVATYFAEALARRIYRLSPQT 267
Db 61 VHALMACAEAIQONNLTAEALVKQIGFLAVSQAGMRKVATYFAEALARRIYRLSPQN 120

QY 268 QIDHSLSDTLQMHFYETCPYLKFAHFTANQAILAEAFEGKKRVHVIDFSMNQGLQWPALMQ 327
Db 121 QIDHCLSDTLQMHFYETCPYLKFAHFTANQAILAEAFEGKKRVHVIDFSMNQGLQWPALMQ 180

QY 328 ALALREGGPPSRLTGTIGPPAADNSDHLHEVGCKLAQLAEAIHVEFYRGFVANSIADLD 387
Db 181 ALALREGGPPPTFRLTGTIGPPAADNSDHLHEVGCKLAQLAEAIHVEFYRGFVANSIADLD 240

QY 388 ASMLELRPSDEAVAVNSVFELHKLGRGTGGIEKVF 424
Db 241 ASMLELRPSDTEAVAVNSVFELHKLGRXGGIEKVLG 277

```

```

RESULT 10
US-08-842-445-34
; Sequence 34, Application US/08842445A
; Patent No. 6441270
; GENERAL INFORMATION:
; APPLICANT: Benfey et al.
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses
; FILE REFERENCE: 5914-056-999
; CURRENT APPLICATION NUMBER: US/08/842,445A
; EARLIER FILING DATE: 1997-04-24
; EARLIER FILING DATE: 1996-04-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Plant
; NAME/KEY: VARIANT
; LOCATION: (1)...(277)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-08-842-445-34

Query Match
Best Local Similarity 41.4%; Score 1222.5; DB 4; Length 277;
Matches 246; Conservative 6; Mismatches 14; Indels 11; Gaps 2;

QY 159 NKRLKSPSSPDSMTSPSP-----AGVIGTIVTTVT-----ESTRPLILVDSQDNGVRL 207
Db 1 NKRLKSCSPDSMTSTGTQIGGVIGTIVTTTITTTTAAAESTRSVILVDSQENGVR 60

QY 208 VHALMACAEAVQSSNLTAEALVKQIGFLAVSQAGMRKVATYFAEALARRIYRLSPQT 267
Db 61 VHALMACAEAIQONNLTAEALVKQIGFLAVSQAGMRKVATYFAEALARRIYRLSPQN 120

QY 268 QIDHSLSDTLQMHFYETCPYLKFAHFTANQAILAEAFEGKKRVHVIDFSMNQGLQWPALMQ 327

```

```

Db 121 QIDHCLSDTLQMHFYETCPYLKFAHFTANQAILAEAFEGKKRVHVIDFSMNQGLQWPALMQ 180
QY 328 ALALREGGPPSRLTGTIGPPAADNSDHLHEVGCKLAQLAEAIHVEFYRGFVANSIADLD 387
Db 181 ALALREGGPPPTFRLTGTIGPPAADNSDHLHEVGCKLAQLAEAIHVEFYRGFVANSIADLD 240
QY 388 ASMLELRPSDEAVAVNSVFELHKLGRGTGGIEKVF 424
Db 241 ASMLELRPSDTEAVAVNSVFELHKLGRXGGIEKVLG 277

```

```

RESULT 11
US-09-186-188B-34
; Sequence 34, Application US/09186188B
; Patent No. 6455672
; GENERAL INFORMATION:
; APPLICANT: Benfey et al.
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses
; FILE REFERENCE: 5914-074-999
; CURRENT APPLICATION NUMBER: US/09/186,188B
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 08/842,445
; PRIOR FILING DATE: 1997-04-24
; PRIOR APPLICATION NUMBER: 08/638,617
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Plant
; NAME/KEY: VARIANT
; LOCATION: (1)...(277)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-186-188B-34

```

```

Query Match
Best Local Similarity 41.4%; Score 1222.5; DB 4; Length 277;
Matches 246; Conservative 6; Mismatches 14; Indels 11; Gaps 2;

QY 159 NKRLKSPSSPDSMTSPSP-----AGVIGTIVTTVT-----ESTRPLILVDSQDNGVRL 207
Db 1 NKRLKSCSPDSMTSTGTQIGGVIGTIVTTTITTTTAAAESTRSVILVDSQENGVR 60

QY 208 VHALMACAEAVQSSNLTAEALVKQIGFLAVSQAGMRKVATYFAEALARRIYRLSPQT 267
Db 61 VHALMACAEAIQONNLTAEALVKQIGFLAVSQAGMRKVATYFAEALARRIYRLSPQN 120

QY 268 QIDHSLSDTLQMHFYETCPYLKFAHFTANQAILAEAFEGKKRVHVIDFSMNQGLQWPALMQ 327
Db 121 QIDHCLSDTLQMHFYETCPYLKFAHFTANQAILAEAFEGKKRVHVIDFSMNQGLQWPALMQ 180

QY 328 ALALREGGPPSRLTGTIGPPAADNSDHLHEVGCKLAQLAEAIHVEFYRGFVANSIADLD 387
Db 181 ALALREGGPPPTFRLTGTIGPPAADNSDHLHEVGCKLAQLAEAIHVEFYRGFVANSIADLD 240

QY 388 ASMLELRPSDEAVAVNSVFELHKLGRGTGGIEKVF 424
Db 241 ASMLELRPSDTEAVAVNSVFELHKLGRXGGIEKVLG 277

```

```

RESULT 12
US-09-186-276B-35
; Sequence 35, Application US/09186276B
; Patent No. 6388173
; GENERAL INFORMATION:
; APPLICANT: Benfey, Philip
; APPLICANT: DiLaurenzio, Laura
; APPLICANT: Wysocka-Diller, Joanna
; APPLICANT: Malamy, Jocelyn E.
; APPLICANT: Pysh, Leonard

```





```
; FILE REFERENCE: 620-45
; CURRENT APPLICATION NUMBER: US/09/117,853
; CURRENT FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: PCT/GB97/00390
; EARLIER FILING DATE: 1997-02-12
; EARLIER APPLICATION NUMBER: GB 9602796.6
; EARLIER FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-117-853-8

Query Match      29.7%; Score 877; DB 3; Length 282;
Best Local Similarity 59.2%; Pred. No. 5.8e-75;
Matches 196; Conservative 23; Mismatches 34; Indels 78; Gaps 9;

QY 1 MKRDLHQFGQPNHGTSTSPAVFGKDKMMVKEEEDDELLGVLGYKVRSEMAEV 60
Db 1 MKRDHHH-----HH-----QDKKTMMNEEDD-----GNGMDV 28

QY 61 ALKLELTMMGNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELNPAAATGSNALNPE 120
Db 29 AOKLEQLEVMMSNVQEDDLSQLATETVHYNPAELYTLWLDMLTDLNPSS-----NAE 81

QY 121 INNNNNSFFTGGDLKAIPGNVCRRSNOFAPAVDSS-----NKRKFPSSSPD 169
Db 82 Y-----DLKAIPGDAIL---NQFAIDSASSSQGGDTYTNRLKCSN--- 123

QY 170 SMWTSPPAGVIGTGVTTVTESTRPLILVDSQONGVRLVHALMACAEAVQSSNLTAEAL 229
Db 124 -----GVVETT-TATAESTRHWLVDSQENGRLVHALLACAEAVQENLTVAEAL 173

QY 230 VKQIGFLAVSOAGAMRKVATYFAEALARRIYRLSPPTQIDHSLSDTLOMHFYETCPYLK 289
Db 174 VKQIGFLAVSQIGAMRKVATYFAEALARRIYRLSPSPIDHSLSDTLOMHFYETCPYLK 233

QY 290 FAHFTANQAILEAFEGKKRVHVDFSMNQGL 320
Db 234 FAHFTANQAILEAFQGGKKRVHVDFSMQGL 264
```

Search completed: November 1, 2004, 21:15:08  
Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 21:13:46 ; Search time 131 Seconds  
(without alignments)  
1415.660 Million cell updates/sec

Title: US-10-030-194A-4

Perfect score: 2950

Sequence: 1 MKRDLHFQGNHGTSTAGS.....LSWHTRLPTTSANKLSAVH 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	2408.5	81.6	587	15	US-10-412-699B-216
2	2110.5	71.5	532	9	US-09-911-513-2
3	2110.5	71.5	532	10	US-09-911-514-2
4	2106.5	71.4	532	14	US-10-278-536-46
5	2106.5	71.4	532	14	US-10-225-066A-784
6	2106.5	71.4	532	15	US-10-374-780A-2290
7	2106.5	71.4	532	15	US-10-412-699B-218
8	1791	60.7	523	15	US-10-424-599-151286
9	1774.5	60.2	544	15	US-10-425-114-42064
10	1748	59.3	381	15	US-10-425-114-38181
11	1748	59.3	595	15	US-10-424-599-268568
12	1594	54.0	630	14	US-10-310-154-569
13	1592	54.0	625	14	US-10-310-154-562

14	1551.5	52.6	609	16	US-10-655-799-40
15	1458.5	49.4	425	15	US-10-425-114-44761
16	1448.5	49.1	546	15	US-10-425-114-72866
17	1350.5	45.8	327	15	US-10-425-114-37141
18	1222.5	41.4	277	9	US-09-186-276B-34
19	1222.5	41.4	277	9	US-09-186-188B-34
20	1222.5	41.4	277	14	US-10-253-007-34
21	1162.5	39.4	508	15	US-10-425-114-55451
22	1162.5	39.4	524	15	US-10-424-599-220386
23	1142	38.7	490	15	US-10-425-114-38253
24	1141	38.7	488	15	US-10-424-599-221067
25	1098	37.2	319	15	US-10-425-114-37774
26	1089.5	36.9	262	9	US-09-186-276B-35
27	1089.5	36.9	262	9	US-09-186-188B-35
28	1089.5	36.9	262	14	US-10-253-007-35
29	1050.5	35.6	399	15	US-10-425-114-37084
30	940	31.9	298	16	US-10-655-799-28
31	937.5	31.8	493	16	US-10-437-963-126535
32	877	29.7	282	9	US-09-911-513-8
33	877	29.7	282	10	US-09-911-514-8
34	846	28.7	259	9	US-09-911-513-6
35	846	28.7	259	10	US-09-911-514-6
36	785	26.6	276	15	US-10-425-114-40354
37	653	22.1	221	9	US-09-911-513-4
38	653	22.1	221	10	US-09-911-514-4
39	632	21.4	532	16	US-10-437-963-190977
40	608.5	20.6	482	16	US-10-437-963-174156
41	597	20.2	542	14	US-10-310-154-568
42	591	20.0	473	15	US-10-424-599-180975
43	588	19.9	460	15	US-10-425-114-38285
44	586	19.9	482	14	US-10-225-066A-220
45	586	19.9	482	14	US-10-225-066A-220

ALIGNMENTS

RESULT 1

- US-10-412-699B-216
- Sequence 216, Application US/10412699B
- Publication No. US20040045049A1
- GENERAL INFORMATION:
- APPLICANT: Mendel Biotechnology, Inc.
- APPLICANT: Zhang, James
- APPLICANT: Fromm, Michael E.
- APPLICANT: Heard, Jacqueline E.
- APPLICANT: Riechmann, Jose Luis
- APPLICANT: Adam, Luc J.
- APPLICANT: Broun, Pierre E.
- APPLICANT: Pineda, Omaira
- APPLICANT: Reuber, T. Lynne
- APPLICANT: Keddle, James S.
- APPLICANT: Yu, Guo-Liang
- APPLICANT: Jiang, Cai-Zhong
- APPLICANT: Samaha, Raymond R.
- APPLICANT: Pilgrim, Marsha L.
- APPLICANT: Creelman, Robert A.
- APPLICANT: DuBell, Arnold N.
- APPLICANT: Ratcliffe, Oliver
- APPLICANT: Kumimoto, Roderick
- APPLICANT: Sherman, Bradley K.
- TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
- FILE REFERENCE: MBI-0048CIP
- CURRENT APPLICATION NUMBER: US/10/412,699B
- CURRENT FILING DATE: 2003-04-10
- PRIOR APPLICATION NUMBER: 09/394,519
- PRIOR FILING DATE: 1999-09-13
- PRIOR APPLICATION NUMBER: 09/489,376
- PRIOR FILING DATE: 2000-01-21
- PRIOR APPLICATION NUMBER: 09/506,720
- PRIOR FILING DATE: 2000-02-17
- PRIOR APPLICATION NUMBER: 09/533,030
- PRIOR FILING DATE: 2000-03-22

```
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 216
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G307
US-10-412-699B-216

Query Match      81.6%; Score 2408.5; DB 15; Length 587;
Best Local Similarity 81.9%; Pred. No. 8e-200;
Matches 490; Conservative 23; Mismatches 48; Indels 37; Gaps 10;

Qy 1 MKRDHQFG--PNHGTISAGSTSSPAVFGKDKMMVKKEE-----DDELLGVLYGVYK 52
Db 1 MKRDHQFGRLSNHGTSSSSSS-----KDKMMVKKEDEGGNMDDELLAVLYGVYK 54

Qy 53 RSSEMAEVALKEQLETVMGNAQEDGLAHLATDTVHYNPAELYSWLDNMLTLPNPPAAT 112
Db 55 RSSEMAEVALKEQLETVMGNSVQEDGLSHLATDTVHYNPSELYSWLDNMLSELNPPPLPA 114

Qy 113 GSNALNPINNNSNP-FTGDDLKAIPGNVCRRSNQFAFVDS-----NKRLLKPS 166
Db 115 SSGGLDPLPSPEICGFPASDYDLKVPGNAI----YQFP-AIDSSSSNNQNKRLKSCS 169

Qy 167 SPDSMVTSPSP-----AGVIGTTT---ESTRPLILVDSQNGVRLVHALMACA 215
Db 170 SPDSMVTSTGTQIGVGITVT-----TAAESTRSVILVDSQNGVRLVHALMACA 229

Qy 216 EAVQSSNLTAEALVKQIGFLAVSQAGAMRKVATYFAEALARRIYRLSPQTQIDHSLSD 275
Db 230 EAIQNNLTAEALVKQIGFLAVSQAGAMRKVATYFAEALARRIYRLSPQNDHCLSD 289

Qy 276 TLOMHFYETCPYLKPAHFTANQAILEAFEGKRVHVIDFSMNQGLQWPAALALREGG 335
Db 290 TLOMHFYETCPYLKPAHFTANQAILEAFEGKRVHVIDFSMNQGLQWPAALALREGG 349

Qy 336 PPSFRLTGI GPPADNSDHLHEVGCKLAQAEAIHVEFEYRGFVANSIADLDASMLELRP 395
Db 350 PPTFRLTGI GPPADNSDHLHEVGCKLAQAEAIHVEFEYRGFVANSIADLDASMLELRP 409

Qy 396 SETEAVAVNSVFLHKLGRGTGGIEKVFVGVVKQIKPVIFTVVEQSNHNGVPFLDRFTES 455
Db 410 SDTEAVAVNSVFLHKLGRGIEKVLGVVKQIKPVIFTVVEQSNHNGVPFLDRFTES 469

Qy 456 LHYSTLTDFSLEGAPSSQDKMSEVYLGKQICNLVACGPDPRVERHETLSQWSNRFSGG 515
Db 470 LHYSTLTDFSLEGVPSQDKMSEVYLGKQICNLVACGPDPRVERHETLSQWGNRFSGG 529

Qy 516 FAPAHLGSAFNAKQASTLLALFNGGEGYRVEKNNGCLMLSWHTRPLITTSAMKLS-AVH 572
Db 530 LAPAHLGSAFNAKQASMLLSVFNSSGGYRVEESNGCLMLGWHTRPLITTSAMKLSAAH 587
```

RESULT 2  
US-09-911-513-2  
; Sequence 2, Application US/09911513  
; Patent No. US2002004995A1  
; GENERAL INFORMATION:

```
; APPLICANT: Harberd, Nicholas P
; APPLICANT: Peng, Jinrong
; APPLICANT: Carol, Pierre
; APPLICANT: Richards, Donald E
; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
; FILE REFERENCE: 620-157
; CURRENT APPLICATION NUMBER: US/09/911,513
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 09/117,853
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: PCT/GB97/00390
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9602796.6
; PRIOR FILING DATE: 1996-02-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-911-513-2

Query Match      71.5%; Score 2110.5; DB 9; Length 532;
Best Local Similarity 73.5%; Pred. No. 5e-174;
Matches 429; Conservative 35; Mismatches 51; Indels 69; Gaps 9;

Qy 1 MKRDHQFGQGNHGTISAGSTSSPAVFGKDKMMVKKEE-----DELLGVLYGVKRSSE 56
Db 1 MKRDHH-----HHQDK-----KTMMEEDDGNMDELLAVLYGVKRSSE 41

Qy 57 MAEVALKEQLETVMGNAQEDGLAHLATDTVHYNPAELYSWLDNMLTLPNPPAATGSA 116
Db 42 MADVAQKLEQLEVMNSVQEDDLSQLATETVHYNPAELYTWLDNMLTDLNPPSS----- 95

Qy 117 LNPEINNNSNPFTGGDLKAIPGNVCRRSNQFAFVDS-----NKRLLKPS 165
Db 96 -NAEY-----DLKAIPGDAIL---NQPAIDSASSNSQGGGDTVTNKRLLKCS 139

Qy 166 SPDSMVTSPSPAGVIGTTT---ESTRPLILVDSQNGVRLVHALMACAEAVQSSNLT 225
Db 140 N-----GVVETT-TATAESTRHVVLDVDSQNGVRLVHALMACAEAVQKENLTV 186

Qy 226 AEALVKQIGFLAVSQAGAMRKVATYFAEALARRIYRLSPQTQIDHSLSDTLQMHFYETC 285
Db 187 AEALVKQIGFLAVSQIGAMRKVATYFAEALARRIYRLSPQSDHSLSDTLQMHFYETC 246

Qy 286 PYLKFHFTANQAILEAFEGKRVHVIDFSMNQGLQWPAALALREGGPPSFLTGIG 345
Db 247 PYLKFHFTANQAILEAFQGGKRVHVIDFSNSQGLQWPAALALRPGGPPVFLTGIG 306

Qy 346 PPAADNSDHLHEVGCKLAQAEAIHVEFEYRGFVANSIADLDASMLELRPSETAEAVNS 405
Db 307 PPAADNFYDLHEVGCKLAHLAEAIHVEFEYRGFVANTLADLDASMLELRPSEIESVANS 366

Qy 406 VFELHKLGRGTGGIEKVFVGVVKQIKPVIFTVVEQSNHNGVPFLDRFTESLHYSTLTDFS 465
Db 367 VFELHKLGRGAIDKVLGVVNVQIKPEIFTVVEQSNHNSPIFLDRFTESLHYSTLTDFS 426

Qy 466 LEGAPSSQDKMSEVYLGKQICNLVACGPDPRVERHETLSQWSNRFSGGFPAPAHLSNA 525
Db 427 LEGVPSQDKMSEVYLGKQICNVVACDGPDRVERHETLSQWNRNRFSGAGPAAAHISNA 486

Qy 526 FKQASTLLALFNGGEGYRVEKNNGCLMLSWHTRPLITTSAMKLS 569
Db 487 FKQASMLLALFNGGEGYRVEESDCLMLGWHTRPLIATSAWKLS 530
```

RESULT 3  
US-09-911-514-2  
; Sequence 2, Application US/09911514  
; Publication No. US20030084470A1  
; GENERAL INFORMATION:  
; APPLICANT: Harberd, Nicholas P

; APPLICANT: Peng, Jinrong  
 ; APPLICANT: Carol, Pierre  
 ; APPLICANT: Richards, Donald E  
 ; TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana  
 ; FILE REFERENCE: 620-158  
 ; CURRENT APPLICATION NUMBER: US/09/911,514  
 ; CURRENT FILING DATE: 2001-07-25  
 ; PRIOR APPLICATION NUMBER: US 09/117,853  
 ; PRIOR FILING DATE: 1998-08-12  
 ; PRIOR APPLICATION NUMBER: PCT/GB97/00390  
 ; PRIOR FILING DATE: 1997-02-12  
 ; PRIOR APPLICATION NUMBER: GB 9602796.6  
 ; PRIOR FILING DATE: 1996-02-12  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 532  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-911-514-2

Query Match 71.5%; Score 2110.5; DB 10; Length 532;  
 Best Local Similarity 73.5%; Pred. No. 5e-174;  
 Matches 429; Conservative 35; Mismatches 51; Indels 69; Gaps 9;

Qy	1	MKRDHQPQGNHGTSTAGSSTSPAVFGKDKMMVKEED----	DELLGVLYKVRSE	56
Db	1	MKRDHGH-----HHQDK-----	KTWMMNEEDDGNMDELLAVLYKVRSE	41
Qy	57	MAEVALKLEQETWGMNAQDGLAHLATDTVHYNPAELYSLDNLMLTELNPAAATTGNSA	116	
Db	42	MADVAQKLEQLEVMNSVQEDDLSQLATETVHYNPAELYTLWDSMLTDLNPPSS-----	95	
Qy	117	LNPEINNNNNNSFFTGGDLKAI PGNAVCRNRNQAFADVSSS-----	NKRLKPS	165
Db	96	-NABY-----	DLKAI PGDAIL--NQFAIDSASSNQGGGDTVTNNKRLKCS	139
Qy	166	SSPDSMTVSPSPAGVIGTTVTTESTRPLILVDSQDNGVRLVHALMACAEAVQSSNLT	225	
Db	140	N-----	GWETT-TATAESTRHVVLDVDSQENGVRVHALLACAEAVQKENLTV	186
Qy	226	AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC	285	
Db	187	AEALVKQIGFLAVSQIGAMRKVATYFABALARRIYRLSPSPIDHSLSDTLQMHFYETC	246	
Qy	286	PYLKFAHTANQAILEAFEGKRVHVHIDFMSQGLQWPAALQALALREGGPPSRLTGIG	345	
Db	247	PYLKFAHTANQAILEAFQGGKRVHVHIDFMSQGLQWPAALQALALRPGGPPVRLTGIG	306	
Qy	346	PPAADNSDHLHEVGCKLAQLAEAIHVEPEYRGFVANSIADLDASMLLRPSETEAVVNS	405	
Db	307	PPAPDNFDYLHEVGCKLAHLAEAIHVEPEYRGFVANTLADLDASMLLRPSEIEVAVNS	366	
Qy	406	VFELHKLGRGTGGIEKVGKQIKPVITFTVVEQSNHNGPVFLDRFTESLHYSTLPDS	465	
Db	367	VFELHKLGRPCAIDKVLGVVQIKPEIFTVVEQSNHNSPIFLDRFTESLHYSTLPDS	426	
Qy	466	LEGAPSSQDKVMSEVYLGKQICNLVACGPDVRVERHETLSQWNRFGSGGAPAPHLGNSA	525	
Db	427	LEGVPSSQDKVMSEVYLGKQICNVVACDGPDRVERHETLSQWNRFGSAGFAAAHIGNSA	486	
Qy	526	FKQASTLIALFNGGEGYRVEKNNGCLMLSWHTRPLITTSAAKLS	569	
Db	487	FKQASMLIALFNGGEGYRVEESDGLMLGWHTRPLIATSAAKLS	530	

RESULT 4  
 US-10-278-536-46  
 ; Sequence 46, Application US/10278536  
 ; Publication No. US2003013386A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Samaha, Raymond  
 ; APPLICANT: Heard, Jacqueline

; APPLICANT: Jiang, Cai-Zhong  
 ; APPLICANT: Pineda, Omaira  
 ; APPLICANT: Reuber, Lynne  
 ; APPLICANT: Riechmann, Jose-Luis  
 ; APPLICANT: Yu, Guo-Liang  
 ; APPLICANT: Keddle, James  
 ; APPLICANT: Ratcliffe, Oliver  
 ; APPLICANT: Pilgrim, Marsha  
 ; APPLICANT: Adam, Luc  
 ; APPLICANT: Broun, Pierre  
 ; TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES  
 ; FILE REFERENCE: MBI-011  
 ; CURRENT APPLICATION NUMBER: US/10/278,536  
 ; CURRENT FILING DATE: 2002-10-22  
 ; PRIOR APPLICATION NUMBER: 60/125,814  
 ; PRIOR FILING DATE: 1999-03-23  
 ; NUMBER OF SEQ ID NOS: 238  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 46  
 ; LENGTH: 532  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; OTHER INFORMATION: G308  
 ; US-10-278-536-46

Query Match 71.4%; Score 2106.5; DB 14; Length 532;  
 Best Local Similarity 73.3%; Pred. No. 1.1e-173;  
 Matches 428; Conservative 36; Mismatches 51; Indels 69; Gaps 9;

Qy	1	MKRDHQPQGNHGTSTAGSSTSPAVFGKDKMMVKEED----	DELLGVLYKVRSE	56
Db	1	MKRDHGH-----HHQDK-----	KTWMMNEEDDGNMDELLAVLYKVRSE	41
Qy	57	MAEVALKLEQETWGMNAQDGLAHLATDTVHYNPAELYSLDNLMLTELNPAAATTGNSA	116	
Db	42	MADVAQKLEQLEVMNSVQEDDLSQLATETVHYNPAELYTLWDSMLTDLNPPSS-----	95	
Qy	117	LNPEINNNNNNSFFTGGDLKAI PGNAVCRNRNQAFADVSSS-----	NKRLKPS	165
Db	96	-NABY-----	DLKAI PGDAIL--NQFAIDSASSNQGGGDTVTNNKRLKCS	139
Qy	166	SSPDSMTVSPSPAGVIGTTVTTESTRPLILVDSQDNGVRLVHALMACAEAVQSSNLT	225	
Db	140	N-----	GWETT-TATAESTRHVVLDVDSQENGVRVHALLACAEAVQKENLTV	186
Qy	226	AEALVKQIGFLAVSQAGAMRKVATYFABALARRIYRLSPPTQIDHSLSDTLQMHFYETC	285	
Db	187	AEALVKQIGFLAVSQIGAMRKVATYFABALARRIYRLSPSPIDHSLSDTLQMHFYETC	246	
Qy	286	PYLKFAHTANQAILEAFEGKRVHVHIDFMSQGLQWPAALQALALREGGPPSRLTGIG	345	
Db	247	PYLKFAHTANQAILEAFQGGKRVHVHIDFMSQGLQWPAALQALALRPGGPPVRLTGIG	306	
Qy	346	PPAADNSDHLHEVGCKLAQLAEAIHVEPEYRGFVANSIADLDASMLLRPSETEAVVNS	405	
Db	307	PPAPDNFDYLHEVGCKLAHLAEAIHVEPEYRGFVANTLADLDASMLLRPSEIEVAVNS	366	
Qy	406	VFELHKLGRGTGGIEKVGKQIKPVITFTVVEQSNHNGPVFLDRFTESLHYSTLPDS	465	
Db	367	VFELHKLGRPCAIDKVLGVVQIKPEIFTVVEQSNHNSPIFLDRFTESLHYSTLPDS	426	
Qy	466	LEGAPSSQDKVMSEVYLGKQICNLVACGPDVRVERHETLSQWNRFGSGGAPAPHLGNSA	525	
Db	427	LEGVPSSQDKVMSEVYLGKQICNVVACDGPDRVERHETLSQWNRFGSAGFAAAHIGNSA	486	
Qy	526	FKQASTLIALFNGGEGYRVEKNNGCLMLSWHTRPLITTSAAKLS	569	
Db	487	FKQASMLIALFNGGEGYRVEESDGLMLGWHTRPLIATSAAKLS	530	

RESULT 5  
 US-10-225-066A-784

; Sequence 784, Application US/10225066A  
; Publication No. US20030226173A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendel Biotechnology, Inc.  
; APPLICANT: RATCLIFFE, Oliver  
; APPLICANT: RIECHMANN, Jose Luis  
; APPLICANT: ADAM, Luc J  
; APPLICANT: DUBELL, Arnold T  
; APPLICANT: HEARD, Jacqueline E  
; APPLICANT: PILGRIM, Marsha L  
; APPLICANT: JIANG, Cai-Zhong  
; APPLICANT: REUBER, T. Lynne  
; APPLICANT: CREELMAN, Robert A  
; APPLICANT: PINEDA, Omaira  
; APPLICANT: YU, Guo-Liang  
; APPLICANT: BROUN, Pierre E  
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants  
; FILE REFERENCE: M010036-2 US  
; CURRENT APPLICATION NUMBER: US/10/225,066A  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/310,847  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 60/336,049  
; PRIOR FILING DATE: 2001-12-05  
; PRIOR APPLICATION NUMBER: 60/338,692  
; PRIOR FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: 10/171,468  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 1122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 784  
; LENGTH: 532  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-225-066A-784

Query Match 71.4%; Score 2106.5; DB 14; Length 532;  
Best Local Similarity 73.3%; Pred. No. 1.1e-173;  
Matches 428; Conservative 36; Mismatches 51; Indels 69; Gaps 9;

QY 1 MKRDLHQFGPNHGTSIAGSSTSPAVFGKDKMMVMKVEED-----DELLGVLYKVRSS 56  
DB 1 MKRDHHH-----HHQDK-----KTMWNEEDDNGMDLAVLYKVRSS 41

QY 57 MAEVALKLEQLETWGMNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELNPAAATTGSNA 116  
DB 42 MADVAQKLEQLEVMNSVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSS-----95

QY 117 LNPEINNNNNNSFTGGDLKATPGNAVCERSNOFAFVDS-----NKLKPS 165  
DB 96 -NAEY-----DLKALPGDAIL---NQFAIDSASSNQGGGDTYTTNKLKCS 139

QY 166 SSPDSMTSPSPAGVIGTIVTTVTBSTPLIIVDSQDNGVRLVHALMACAFVQSSNLT 225  
DB 140 N-----GVVEIT-TATAESTRHVLDVSQENGVLVALLACAEVQENLT 186

QY 226 AEALVKQIGFLAVSAGAMRVATYFABALARRIYRSPPTQIDHSLSDTLQMHFYETC 285  
DB 187 AEALVKQIGFLAVSAGAMRVATYFABALARRIYRSPSPIDHSLSDTLQMHFYETC 246

QY 286 PYLKFAHTANQAILEAFEGKRVHVHIDFSMNQGLQWALMOALREGGPPSFRLTGIG 345  
DB 247 PYLKFAHTANQAILEAFQKRVHVHIDFSMQGLQWALMOALRPGGPPVFRLTGIG 306

QY 346 PPAADNSHLHEVGKLAQLAEAIVHFEYRGFVANSADLADSMLELRPSETEAVAVNS 405  
DB 307 PPAADNSHLHEVGKLAQLAEAIVHFEYRGFVANSADLADSMLELRPSETEAVAVNS 366

QY 406 VFELHKLGRGTGTEKRVGVKQIKPVIFTVVEQSNHNGPFVLDRTFESLHYITLFD 465  
DB 367 VFELHKLGRGCAIDKVLGVVQIKPEIFTVVEQSNHNSPIFLDRFTFESLHYITLFD 426

QY 466 LEGAPSSQDKVMSEVYLKQICNLVACGPDVRVERHETLSQWNRFGSGFAPAHLSNA 525  
DB 427 LEGVPSGQDKVMSEVYLKQICNVVACDGPDRVERHETLSQWNRFGSAGFAAAHGSNA 486

QY 526 FKQASTILALFNGGEGYRVERKNGCLMLSWHTRPLITTSNAKLS 569  
DB 487 FKQASMLLALFNGGEGYRVEESDGLMLGWHTRPLIATSAKLS 530

## RESULT 6

US-10-374-780A-2290  
; Sequence 2290, Application US/10374780A  
; Publication No. US20040019927A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, Bradley K  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Heard, Jacqueline E  
; APPLICANT: Haake, Volker  
; APPLICANT: Creelman, Robert A  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Adam, Luc J  
; APPLICANT: Reuber, T. Lynne  
; APPLICANT: Keddies, James E  
; APPLICANT: Broun, Pierre E  
; APPLICANT: Pilgrim, Marsha L  
; APPLICANT: Dubell III, Arnold T  
; APPLICANT: Pineda, Omaira  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
; FILE REFERENCE: M01-0047 CIP  
; CURRENT APPLICATION NUMBER: US/10/374,780A  
; CURRENT FILING DATE: 2003-02-25  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/310,847  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 09/934,455  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/336,049  
; PRIOR FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: 60/338,692  
; PRIOR FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: 10/171,468  
; PRIOR FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 10/225,066  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 10/225,067  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 10/225,068  
; NUMBER OF SEQ ID NOS: 2906  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2290  
; LENGTH: 532  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: G308  
US-10-374-780A-2290

Query Match 71.4%; Score 2106.5; DB 15; Length 532;  
Best Local Similarity 73.3%; Pred. No. 1.1e-173;  
Matches 428; Conservative 36; Mismatches 51; Indels 69; Gaps 9;

QY 1 MKRDLHQFGPNHGTSIAGSSTSPAVFGKDKMMVMKVEED-----DELLGVLYKVRSS 56  
DB 1 MKRDHHH-----HHQDK-----KTMWNEEDDNGMDLAVLYKVRSS 41

QY 57 MAEVALKLEQLETWGMNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELNPAAATTGSNA 116  
DB 42 MADVAQKLEQLEVMNSVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSS-----95



; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 151286  
; LENGTH: 523  
; TYPE: PRT  
; ORGANISM: Glycine max  
; NAME/KEY: unsure  
; LOCATION: (1)..(523)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_107634C.1.pep  
US-10-424-599-151286

Query Match 60.7%; Score 1791; DB 15; Length 523;  
Best Local Similarity 64.5%; Pred. No. 2.6e-146;  
Matches 369; Conservative 59; Mismatches 84; Indels 60; Gaps 13;

QY 1 MKRDLHQFGPNHGTSIAGSSSPAVFGKDKMMVKKEED--DELLGVLGKVRSSMA 58  
DB 1 MKRERQQL-GSNAGTSSCGYSS-----GKSNLW---EEEGMDLLEAVVGYKVRSSDMA 50  
QY 59 EVALKLEQLETMMGNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELNPAAATTGSNALN 118  
DB 51 EVAQKLERLEAENGVDQD-LTDLNDVAVHNPDSISNWLQTMLSNFDPLPSE-----E 103  
QY 119 PEINNNNNFTGGDLKAIPGNVAVRRSNQFAFVAVDSSNNKRLKPSPPSMTVSPSPA 178  
DB 104 PE-----KDSASSDYDLKAIPGKAIY-----GAGSDAL---PNPK 135  
QY 179 GVIGTIVTTVTESTRPLILVDSQDNGVRLVHALMACAEAVQSSNLTILAEALVKQIGFLAV 238  
DB 136 RV-----RADSRRAVVVDSENGIRLVHSLMACAEAVENNLLAVAEALVKQIGFLAV 189  
QY 239 SQAGAMRKVATYFAEALARRIYRLSPPTQIDHSLSDTLQMHFYETCPYLKFAHFTANQA 298  
DB 190 SQVGARXKATYFAEALARRIYRVFPQ-----HSLSDSLQIHFYETCPYLKFAHFTANQA 245  
QY 299 ILEAFEGKRVHVIDFSMNQGLQWALMQLALREGGPPSFRLTGIGPPAADNSDHLHEV 358  
DB 246 ILEAFQGNRVHVIDFGINQGMQWALMQLALRNDGPPVFRLTGIGPPAADNSDHLQEV 305  
QY 359 GCKLAQLAEAIHVEFEYRGFVANSIADLDASMLELRPSETEAVAVNSVFEHLKLGRTGG 418  
DB 306 GWKLAQLAERIHVQFYRGFVANSIADLDASMLDUR--EDSVAVNSVFEHKLARPGA 363  
QY 419 IEKVFGVVKQIKPVIPTVVEQESNHNHGVFLDRFTESLHYISTLFDLSLEGAP--SSQDKVM 477  
DB 364 VEKLVSVVRQIRPEILTVVEQEAHNHGLSFVDRFTESLHYISTLFDLSBGSVPVNDKAM 423  
QY 478 SEVYLGKQICNLVACGPDVERHETLSQWNRFGSSGFAFAHLGSAFNAFKOASTLLALFN 537  
DB 424 SEVYLGKQICNVVACGMDRVERHETLNQWRNRFSTGFSVHLGSAFNAFKOASTLLSLFG 483  
QY 538 GGEYRVKKNNGCLMLSWHTRPLITTSAAKLS 569  
DB 484 GGDGYRVEENNGCLMLGWHTRPLIATSVWQLA 515

RESULT 9  
US-10-425-114-42064  
; Sequence 42064, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 42064  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Glycine max  
; NAME/KEY: unsure  
; LOCATION: (1)..(544)  
; OTHER INFORMATION: Clone ID: 700738093\_FLI.pep  
US-10-425-114-42064

Query Match 60.2%; Score 1774.5; DB 15; Length 544;  
Best Local Similarity 64.2%; Pred. No. 7.4e-145;  
Matches 368; Conservative 59; Mismatches 85; Indels 61; Gaps 12;

QY 1 MKRDLHQFGPNHGTSIAGSSSPAVFGKDKMMVKKEED--DELLGVLGKVRSSMA 58  
DB 28 MKRERQQL-GSIAGTSSCGYSS-----GKSNLW---EEEGMDLLEAVVGYKVRSSDMA 77  
QY 59 EVALKLEQLETMMGNAQEDGLAHLATDTVHYNPAELYSWLDNMLTELNPAAATTGSNALN 118  
DB 78 EVAQKLERLEAENGVDQD-LPEISNDVAVHNPDSISNWLQTMLSNFDPLPSE-----E 130  
QY 119 PEINNNNNFTGGDLKAIPGNVAVRRSNQFAFVAVDSSNNKRLKPSPPSMTVSPSPA 178  
DB 131 PE-----KDSASSDYDLKAIPGKAIYASDAL-----PNPK 161  
QY 179 GVIGTIVTTVTESTRPLILVDSQDNGVRLVHALMACAEAVQSSNLTILAEALVKQIGFLAV 238  
DB 162 RV-----KADESRRAVVVDSENGIRLVHSLMACAEAVENNLLAVAEALVKQIGFLAV 215  
QY 239 SQAGAMRKVATYFAEALARRIYRLSPPTQIDHSLSDTLQMHFYETCPYLKFAHFTANQA 298  
DB 216 SQVGAMRKVATYFAEALARRIYRVFP-----LQHSLSDSLQIHFYETCPYLKFAHFTANQV 271  
QY 299 ILEAFEGKRVHVIDFSMNQGLQWALMQLALREGGPPSFRLTGIGPPAADNSDHLHEV 358  
DB 272 ILEAFQGNRVHVIDFGINQGMQWALMQLALRNDGPPVFRLTGIGPPAADNSDHLQEV 331  
QY 359 GCKLAQLAEAIHVEFEYRGFVANSIADLDASMLELRPSETEAVAVNSVFEHLKLGRTGG 418  
DB 332 GWKLAQLAERIHVQFYRGFVANSIADLDASMLDUR--EGEAVAVNSVFEHKLARPGA 389  
QY 419 IEKVFGVVKQIKPVIPTVVEQESNHNHGVFLDRFTESLHYISTLFDLSLEGAP--SSQDKVM 477  
DB 390 VEKLVSVVRQIRPEILTVVEQEAHNHGLSFVDRFTESLHYISTLFDLSBGSVPVNDKAM 449  
QY 478 SEVYLGKQICNLVACGPDVERHETLSQWNRFGSSGFAFAHLGSAFNAFKOASTLLALFN 537  
DB 450 SEVYLGKQICNVVACGMDRVERHETLNQWRNRFSTGFSVHLGSAFNAFKOASTLLALFA 509  
QY 538 GGEYRVKKNNGCLMLSWHTRPLITTSAAKLS 570  
DB 510 GGDGYRVEENNGCLMLGWHTRPLIATSAWQLAA 542

RESULT 10  
US-10-425-114-38181  
; Sequence 38181, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With



; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425.114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 38181  
; LENGTH: 381  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB25-063-C7\_FLI.pep  
US-10-425-114-38181

Query Match 59.3%; Score 1748; DB 15; Length 381;  
Best Local Similarity 88.1%; Pred. No. 8.5e-143;  
Matches 334; Conservative 20; Mismatches 25; Indels 0; Gaps 0;  
  
QY 191 STRPLIVDSODNGVRLVHALMACAEAVQSSNLTIAEALVKOIGFLAVSQAGAMRKVATY 250  
DB 1 STRHVIVDSOENGVRLVHALMACAEAVQKENLTVAEALVKOIGFLAVSQIGAMRKVATY 60  
  
QY 251 FAEALARRIYRLSPQPTQIDHSLDYLQMHFYETCPYLKFAHFTANQAILEAFEGKKRVH 310  
DB 61 FAEALARRIYRLSPQSIDHSLDYLQMHFYETCPYLKFAHFTANQAILEAFQKKRVH 120  
  
QY 311 VIDFSMQGLQWPAALMAALAREGPPSFRUTGTGIPPAADNSDHLHEVCGCKLAOLAEAIH 370  
DB 121 VIDFSMQGLQWPAALMAALAREGPPVFRUTGTGIPPAADNFYDLHEVCGCKLAOLAEAIH 180  
  
QY 371 VEFYRGFVANSADLDASMLELRPSETEAVAVNSVFLHKLGRGTGIEKVFVGVQIK 430  
DB 181 VEFYRGFVANTLADLDASMLELRPSEIESVAVNSVFLHKLGRPGAIDKVLGVVQIK 240  
  
QY 431 PVITVVEQSNHNGPVFLDRFTSLHYSTLFDLEGAPSSQDKWSEVYLQICNLV 490  
DB 241 PEITVVEQSNHNSPIFLDRFTSLHYSTLFDLEGVPQDKWSEVYLQICNVV 300  
  
QY 491 ACEGPDVRHERHETLSQWNSRFGSGFAPAHLGSAFQAOSTLLALFNGGEGYRVEKNKGC 550  
DB 301 ACDDPDVRHERHETLSQWNRFGSAGFAAAHTGSAFQAOSTLLALFNGGEGYRVEESDGC 360  
  
QY 551 LMLSWHTRPLITTSNAKLS 569  
DB 361 LMLGWHTRPLIATSNAKLS 379

RESULT 11  
US-10-424-599-268568  
; Sequence 268568, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424.599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 268568  
; LENGTH: 595  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_8453C.1.pep  
US-10-424-599-268568

Query Match 59.3%; Score 1748; DB 15; Length 595;  
Best Local Similarity 60.0%; Pred. No. 1.7e-142;  
Matches 369; Conservative 64; Mismatches 104; Indels 78; Gaps 14;

QY 1 MKRDLHOFQGNHGTSTAGSSTSSPAVFGK-----DKMMVKKEED----- 41  
DB 1 MKRD-----HKDSCGGGAAGTGVKGCSSMQSNGKAKMWEEBEEQQQQQQQQQQ 52  
  
QY 42 --DELLGVLYGKVRSSMAEVALKLEOLETWGNAQEDGLAHLATDTVHYNPAEYLSWLD 99  
DB 53 GMDLLAALGYKVRASDMADVAKLEQLEWMVGCQEDGISHLASDTHVYDPTDLYSWVQ 112  
  
QY 100 NMLTELNPPAATGTSNALNPE---INNNNN--NS---FFTGG---DLKAIPGNVCRSN 148  
DB 113 SMLTELNPEP---NNNLDPSFLIDNNNNIINSTAPVFNDDSEYDLRAIPGIAAPSSL 168  
  
QY 149 QPAFAVD-----SSNKRKLPK---SSPDSMTVTSPPAGVIGTIVTTVTTSTRLPLIVSQ 201  
DB 169 PQDNLHDEIETANNINKRLKPSAESADSASEP-----TRHVVLVDHQ 212  
  
QY 202 DNGVRLVHALMACAEAVQSSNLTIAEALVKOIGFLAVSQAGAMRKVATYFAEALARRIYR 261  
DB 213 EAGVRLVHTLLACAEAVQENLKLADALVKHVGILAAQAGAMRKVASYFAQALARRIYG 272  
  
QY 262 LSPPTQIDHSLDYLQMHFYETCPYLKFAHFTANQAILEAFEGKKRVHVIDFSMQGLQ 321  
DB 273 IFPEET-LDSSFSVLLHMHFYESCPLYLKFAHFTANQAILEAFATAGRVHVIDFGLKQMGQ 331  
  
QY 322 WPALMQALALAREGPPSFRUTGTGIPPAADNSDHLHEVCGCKLAOLAEAIHVEFEYRGFVAN 381  
DB 332 WPALMQALALAREGPPPTFRUTGTGIPPOPDNTDALQQVGWKLQAQIIGVQFEFRGFVCN 391  
  
QY 382 SLADLDASMLELRPSETEAVAVNSVFLHKLGRGTGIEKVFVGVQIKLPVITVVEQES 441  
DB 392 SLADLDPNMLEIRPG--EAVAVNSVFLHRLMARSGVDKVLDTVKINLPQIVITVEQEA 449  
  
QY 442 NHNGFVFLDRFTSLHYSTLFDLEGAPS-----SODKVMSEVYLQICNLVACE 493  
DB 450 NHNGFVFLDRFTSLHYSTLFDLEGSSSSSTGLGSPSQDLLMSELYLGRQICNVVAVE 509  
  
QY 494 GPDVRHERHETLSQWNSRFGSGFAPAHLGSAFQAOSTLLALFNGGEGYRVEKNKGCML 553  
DB 510 GPDVRHERHETLQWRGLDSAGFDPVHLGSAFQAOSTLLALFAGGCGYRVEENNGCLML 569  
  
QY 554 SWHTRPLITTSNAKLS 568  
DB 570 GWHTRPLIATSNAKLS 584

RESULT 12  
US-10-310-154-569  
; Sequence 569, Application US/10310154  
; Publication No. US20030233670A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Chomet, Paul S.  
; APPLICANT: Adams, Thomas H  
; APPLICANT: Ruff, Thomas G.  
; APPLICANT: Agarwal, Ameeta K.  
; APPLICANT: Ahrens, Jeffrey E.  
; APPLICANT: Ball, James A.  
; APPLICANT: Banu, G.  
; APPLICANT: Bell, Erin  
; APPLICANT: Boddupalli, Raghava  
; APPLICANT: Deikman, Jill  
; APPLICANT: Deng, Molian  
; APPLICANT: Dong, Jinzhao  
; APPLICANT: Duff, Stephen M.  
; APPLICANT: Galligan, Meghan M.  
; APPLICANT: Hinchey, Brenda S.  
; APPLICANT: Huang, Shihshieh  
; APPLICANT: Johnson, G. Richard  
; APPLICANT: Jung, Vincent  
; APPLICANT: Kretzmer, Keith A  
; APPLICANT: Laccetti, Lucille B.  
; APPLICANT: Lai, Chao-Qiang  
; APPLICANT: Lee, Gary



Query Match	54.0%	Score 1592;	DB 14;	Length 625;	
Best Local Similarity	52.5%;	Pred. No. 6.3e-129;			
Matches 341; Conservative	67;	Mismatches 137;	Indels 104;	Gaps 14	
Qy	1	MKRD LHQFGQGNHGT	SIAGS	STSSPAVFGKQMM--MVKEBED-DELLGLVGYKVRSE	57
Db	1	MKREYQEA	GGSGGSSADMGSC-----	KDKVMAGAAGEBEDVDELLAALGYKVRSSDM	54
Qy	58	AEVALKULE	QLEETMG-----	NAQEDG-LAHLATDTVHYNPAEYLSYMDNMLTEIN---	106
Db	55	ADVAQKLEQLEWAMG	MGVSPAG	GAADDGFSVSHLATDTVHYNPSDLSWSVMSLSINAPL	114
Qy	107	-----	PPAATG	SNALPEINNNNNNFETGGDLKAJPGNACVRRSQPAFVADSSNNK	160
Db	115	PPIPAPPAARHAST--	SSTVTGGGSGGFF-----	ELP-----	AAADSSSTY 155
Qy	161	RLKPSSSPDS	SVMTGSPAGVI-----	GTVTVTVTES	191
Db	156	ALRPISLPVVATADP	SAADSRDTRKMR	TGGGTSSTSSSSSSLLGGASRGSRVVEAAPPA	215
Qy	192	TR-----	PLILDSDQNGVRLVHALMACAE	AVQSSNLTAEALVKOIGFLAVSQA	241
Db	216	TQGAANAANAPV	VVVVDTSQAGIRLVHAL	CACAEVQOENPAAAEALVKQIPTTLAASQG	275
Qy	242	GAMRKVATYFAEAL	ARRIYRLSP-PQOTIDH	SLSDTLQWHFYETCPYLKFAHFTANQAIL	300
Db	276	GAMRKVAAYFGEAL	ARRYVRFPADSTLLD	AAFADLLHAHFVYESCPYLKFAHFTANQAIL	335
Qy	301	EAPFGKRVHVVD	SRMNOGLQWPA	LMQALAREGGPPSFRLTGIGPPAADNSDHLHVEVC	360
Db	336	EAFAGCHRHVVD	FGIKQGMQWPA	LQALALRPGGPPSFRLTGVPQPQDETALQOQVGW	395
Qy	361	KLAQLAEAIHVEFE	YRGFVANS	LADLDASMLELR-----	PSETEAVAVNSVPFELHKLGR 415
Db	396	KLAQFAHTIRVD	FQYRGLVAATLAD	LEPFMLQPEGEADANEPEVIAVNSVPFELHRLAQ	455
Qy	416	TGGIEKVP	GVVKQIKPVI	FTVVVEQBSNNGPVFLDRFTESLHYSTLFDLSLEGAPSSQ--	473
Db	456	PGALEKVLGT	VHVRPRITVTVVEQ	EANINSGSFLDRFTESLHYSTLFTDLSLEGSSGQAE	515
Qy	474	-----	DKWMSVYLKGQICNL	VACBGPDRVERHETLSQWSNRFGSSGFGAPAHIL	521
Db	516	LSPPAAGGGG	GTQWMSVYLRQICNL	VACBGAERTENHETLQWRNLGRAGFEPVHL	575
Qy	522	GSNAFQKASTLL	ALLFNGGEGYR	VEKNNGCLMLSHTRPLITTSAAKLSA	570
Db	576	GSNAVKQASTLL	ALLPAGDG	YRVEKEGCLTIGWHTRPLIATSARVRAA	624

```

RESULT 14
US-10-6555-799-40
; Sequence 40, Application US/10655799
; Publication No. US20040126843A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Hall, Claire
; APPLICANT: Norriss, Michael Geoffrey
; APPLICANT: Saulsbury, Keith Martin
; TITLE OF INVENTION: Compositions Isolated
; TITLE OF INVENTION: Grasses and Methods
; FILE REFERENCE: 11000.1074U
; CURRENT APPLICATION NUMBER: US/10/655,799
; CURRENT FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/408,782
; PRIOR FILING DATE: 2002-09-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Festuca arundinacea
US-10-6555-799-40

```

Query Match	52.6%	Score 1551.5	DB 16	Length 609
Best Local Similarity	52.7%	Pred. No. 2e-125		
Matches 328	Conservative 147	Mismatches 147	Indels 67	Gaps 14
Qy	1	MKRDLHGFQGNHGTSTAGSGTSPAVFGKDKMM--VKBEED--DELLGVGLGYKVRSE	56	
Db	1			
Db	1	MKRE-YQDAG-----GSSAGDGMGSKDMMSPAPQADEDDVDELLAALGYKVRSSD	51	
Qy	57	MAEVALKLEQLETMG-----NAQDGL-AHLATDTVHYNPAELYSWLDNMLTELNPPAAT	111	
Db	52	MADVAQKLEQLEMAGMGVPAPDDGFTTHIATVHYNPTDLSWVESMSELNAPPPL	111	
Qy	112	TGSNALNPEINNNNNNSFFTGGDLKATPGNACVRRSNOFA-----FAVDSSEN--	159	
Db	112	PPAPRLAPASVTAADGFFD-----IPPPVSDSSSTYALRPIPPSPADLSADLSADSPR	165	
Qy	160	--KRLKP---SSSPDSMVTSPSPAGVIGTIVTTVTESTR-----PLILVSQNGVRLVHAL	211	
Db	166	DPKEMRTGGSTSSSSSSSSSLGCVVEAAPAAAEANAIALPVVVADTQEAGIRLVHAL	225	
Qy	212	MACAEAVQSSNLTIAEALVKQIGFLAVSQAGAMRKVATYFAEALARRVYRL--SPQTOI	269	
Db	226	LACAEAVQENFSAEALVKQI PLLASQGGAMRKVAAVFGAALARRVFRFPQDPSHL	285	
Qy	270	DHLSLSDTLQMHFYETCPYLKFAHFTANQAILLEAFEGKKRVHVIDFSMNQGLQWPAALMAL	329	
Db	286	DAAPADLLHAHFYESSCYLKFAHFTANQAILLEAFAGCRRVHVDFGIKQGMWPAALQAL	345	
Qy	330	ALRGGGPPSFLTGIGPPAADNSDHLHEVGCKLAQAEAIHVEFEYRGFVANSIADLDAS	389	
Db	346	ALRPGGPPSFLTGCGVPPQDETALQOVGWKLQAFHTIGVDFQYRGLVAATLADLSPF	405	
Qy	390	MLELRP-----SETEAVAVNSVFELHKLGLRTGGTEKVFVVVKQIKPVIFTVVEQESNHN	444	
Db	406	MLQPEAEDGPNNEEPIAVNSIFEMHRLLAQPGALEKVLGTVRAVRPRIVTVVEQEAHNH	465	
Qy	445	GPVFLDRFTESLHYSTILFDSLEGAPS-----SQDKMSEVYLKQKICN	488	
Db	466	AGSFDRFTESLHYSTYMFDSLEGAGGSPSEISSGPAAAAAAGPTDQVMSEVYLGRQICN	525	
Qy	489	LVACEGPDPRVERHETLGSWNRRFGSSGPAPAHLSNAFKQASTLILALFNGCGEYVEKNN	548	
Db	526	VVACEGAERTERHETLGHWRGRLGHAGFETVHLSGSNAYKQASTLLALFAGGDGYKVDKE	585	
Qy	549	GCLMLSWHTRPLITTSAAWKLSA	570	
Db	586	GCLTLGWHTRPLIATSARMAA	607	

```

RESULT 15
US-10-425-114-44761
; Sequence 44761, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44761
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701068676

```

US-10-425-114-44761

```

Query Match          49.4%; Score 1458.5; DB 15; Length 425;
Best Local Similarity 68.3%; Pred.No. 1.3e-117;
Matches 287; Conservative 43; Mismatches 61; Indels 29; Gaps 5;

QY 159 NKRLKPS--SSPDSMVTSPSPAGVIGTIVTTTESTRPLILVDSQDNGVRLVHALMACAE 216
Db 14 NKRLKPSPAESADSAASEP-----TRHVVLVDHQEAGVRLVHTLLACAE 57

QY 217 AVQSSNLTAEALVKOIGFLAVSOAGAMRKVATYFAEALARIYRLSPPTOIDHSLSDT 276
Db 58 AVQENLKLADALVKHVGILAAASOAGAMRKVASYFAQALARIYGIFFEET-LDSSFSDV 116

QY 277 LQWHFYETCPYLKFAHFTANQAILFAFECKKRVHVIDPSMNOGLOWPALMQALALREGGP 336
Db 117 LHMHFYESCPLYLKFAHFTANQAILLEAFATAGKVHVIDFGLKQGMOWPALMQALALRPGGP 176

QY 337 PSFRLTGIGPPAADNSDHLHEVGCKLAQAEAIHVEFEYRGFVANSADLADLASMLELRPS 396
Db 177 PTFRLTGIGPPQPDNTDALQQVGLKLAQLAIIQVQFRGFCVNSLADLDPNMLEIRPG 236

QY 397 ETEAVAVNSVFELHKLGHGTGIEKVFGVVKQIKPVIPTVVEQESNHNHGPVFLDRFTESL 456
Db 237 --EAVAVNSVFELHRLARSGSVDKVLDIVKKINPQIVTIVEQEAHNHNGPGFLDRFTEAL 294

QY 457 HYYSTLFDLSLEGAPS-----SODKVMSEVYLGKQICNLVACRGPDPRVERHETLSOWS 508
Db 295 HYYSLFDLSLESGSSSTGLGSPSQDLMSELYLGRQICNVVAYEGDPRVERHETLTQWR 354

QY 509 NTFGSGGFAPAHLGSNAPKQASTLLALFNGGEGYRVEKNNGCLMLSWHTRPLITTSANKL 568
Db 355 GRLDGAGDPVHLGSGNAPKQASMLLALFAGGDGYRVEENNGCLMLGWHTRPLIATSANKL 414
```

Search completed: November 1, 2004, 21:25:31  
Job time : 134 secs